PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



51) International Patent Classification ⁶ :		(11) International Publication Number: WO 00/14106	
C07K 1/00	Al	(43) International Publication Date: 16 March 2000 (16.03.00)	
(21) International Application Number: PCT/(22) International Filing Date: 9 September 1999	US99/210 9 (09.09.9	CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC	
30) Priority Data: 09/150,489 9 September 1998 (09.09) (71) Applicant: THE BURNHAM INSTITUTE [US/US Toriey Pines Road, La Jolla, CA 92037 (US).	Published With international search report. Before the expiration of the time limit for amending th claims and to be republished in the event of the receipt of amendments.		
(72) Inventors: REED, John, C.; 17044 El Camino R Santa Fe, CA 92067 (US). TAKAYAMA, SI Stratford Court #3, Del Mar, CA 92014 (US).	hinichi, 3		
(74) Agents: WONG, James, J. et al.; Campbell & Flore 700, 4370 La Jolla Village Drive, San Diego. (US).	, CA 921		
(54) Title: NOVEL BAG PROTEINS AND NUCLEIO	C ACID N	DIECULES ENCODING THEM	
(57) Abstract			
The present invention provides a family of BAG-BAG-5), the invertebrate <i>C. elegans</i> (BAG-1, BAG-2) that encode them.	–1 related and the fi	proteins from humans (BAG-1L, BAG-1, BAG-2, BAG-3, BAG-4 and ion yeast <i>S. pombe</i> (BAG-1A, BAG-1B) and the nucleic acid molecule	

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovema
AM	Armenia	Fi	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	SZ	Swazifand
AZ	Azerbaijan	GB	United Kingdom	MC	Монасо	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
вв	Barbados	GH	Ghana	MG	Madagascar	TJ	Tapkistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benn	IE	heland	MN	Mongolia	UA	Ukrame
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
$C\Pi$	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
cv	Cuba	KZ	Hazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucra	RU	Russian Federation		
DE	Germany	1.1	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NOVEL BAG PROTEINS AND NUCLEIC ACID MOLECULES ENCODING THEM

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY-SPONSORED RESEARCH AND DEVELOPMENT

This invention was made with government support under grant number CA-67329 awarded by the National Institutes of Health. The United States Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

10 FIELD OF THE INVENTION

This invention relates generally to the fields of molecular biology and molecular medicine and more specifically to a novel family of proteins that can regulate protein folding. The functions of these proteins are potentially diverse, including promoting tumor cell growth and metastasis.

BACKGROUND INFORMATION

The Hsc70/Hsp70-family of molecular chaperones participate in protein folding reactions, controlling 20 protein bioactivity, degradation, complex assembly/disassembly, and translocation across membranes. These proteins interact with hydrophobic regions within target proteins via a carboxyl (C)-terminal peptide binding domain, with substrate binding and release being controlled by the N-terminal ATP-binding domain of Hsc70/Hsp70. Hsc70/Hsp70-assisted folding reactions are accomplished by repeated cycles of peptide binding, refolding, and release,

2

which are coupled to ATP hydrolysis by the ATP-binding domain (ATPase) of Hsc70/Hsp70 and by subsequent nucleotide exchange. The chaperone activity of mammalian Hsc70/Hsp70 is regulated by partner proteins that either modulate the peptide binding cycle or that target the actions of these chaperones to specific proteins and subcellular compartments. DnaJ-family proteins (Hdj-1/Hsp40; Hdj-2; Hdi-3) stimulate the ATPase activity of Hsc70/Hsp70, resulting in the ADP-bound state which binds tightly to peptide substrates. The Hip protein collaborates with Hsc70/Hsp70 and DnaJ homologues in stimulating ATP hydrolysis, and thus also stabilize Hsc70/Hsp70 complexes with substrate polypeptides, whereas the Hop protein may provide co-chaperone functions through interactions with the C-terminal peptide binding domain.

The Bcl-2 associated athanogene-1 (bag-1) is named from the Greek word athanos, which refers to anti-cell death. BAG-1 was previously referred to as Bcl-2-associated protein-1 (BAP-1) in U.S. Patent No. 5,539,094 issued July 23, 1996, which is incorporated herein by reference. In this earlier patent, BAG-1 is described as a portion of the human BAG-1 protein, absent the N-terminal amino acids 1 to 85. In addition, a human protein essentially identical to human BAG-1 was described by Zeiner and Gehring, (Proc. Natl. Acad. Sci., USA 92:11465-11469 (1995)). Subsequent to the issuance of U.S. Patent 5,539,094 the N-terminal amino acid sequence from 1 to 85 of human BAG-1 was reported.

BAG-1 and its longer isoforms BAG-1M (Rap46) and 30 BAG-1L are recently described Hsc70/Hsp70-regulating proteins. BAG-1 competes with Hip for binding to the Hsc70/Hsp70 ATPase domain and promotes substrate release. BAG-1 also reportedly stimulates Hsc70-mediated ATP

10

3

hydrolysis by accelerating ADP/ATP exchange, analogous to the prokaryotic GrpE nucleotide exchange protein of the bacterial Hsc70 homologue, DnaK. Gene transfection studies indicate that BAG-1 proteins can influence a wide variety of cellular phenotypes through their interactions with Hsc70/Hsp70, including increasing resistance to apoptosis, promoting cell proliferation, enhancing tumor cell migration and metastasis, and altering transcriptional activity of steroid hormone receptors.

Despite the notable progress in the art, there remains an unmet need for the further identification and isolation of additional homologous BAG protein species, and the nucleic acid molecules and/or nucleotide sequences that encode them. Such species would provide additional means by which the identity and composition of the BAG domain, that is, the portion of the protein that is influencing or modulating protein folding, could be identified. In addition, such species would be useful for identifying agents that modulate apoptosis as candidates for therapeutic agents, in particular, anticancer agents. The present invention satisfies these need, as well as providing substantial related advantages.

SUMMARY OF THE INVENTION

The present invention provides a family of BAG-1 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO: 4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate C.elegans [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast S.pombe [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)] and the nucleic acid molecules that encode them.

4

Another aspect of the present invention provides an amino acid sequence present in the family of BAG-1 related proteins, that modulates Hsc70/Hsp70 chaperone activity, that is, the BAG domain.

Another aspect of the present invention provides novel polypeptide and nucleic acid compositions and methods useful in modulating Hsc70/Hsp70 chaperone activity.

Another aspect of the present invention is directed to methods for detecting agents that modulate the binding of the BAG family of proteins, such as BAG-1 (beginning at residue 116 of SEQ ID NO:2), and related proteins with the Hsc70/Hsp70 Family of proteins or with other proteins that may interact with the BAG-Family proteins.

Still another aspect of the present invention is directed to methods for detecting agents that induce the dissociation of a bound complex formed by the association of BAG-Family proteins with Hsc70/Hsp70 Family molecule chaperones or other proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the full length cDNA sequence for human BAG-1 (SEQ ID NO:1) protein with the corresponding amino acid sequence (SEQ ID NO:2). Within the full length sequence are included the overlapping sub-sequences of BAG-1 (beginning at nucleotide 391), BAG-1M [beginning at nucleotide 260 of (SEQ ID NO:2)], and BAG-1L [beginning at nucleotide 46 of (SEQ ID NO:2)].

5

Figures 2A and 2B combined shows the full length cDNA sequence (SEQ ID NO:3) aligned with the corresponding amino acid residues for human BAG-2 protein (SEQ ID NO:4).

Figure 3 shows a cDNA sequence (SEQ ID NO:5) saligned with the corresponding amino acid residues for human BAG-3 protein (SEQ ID NO:6).

Figure 4 shows the a cDNA sequence (SEQ ID NO:7) aligned with the corresponding amino acid residues for human BAG-4 protein (SEQ ID NO:8).

Figure 5 shows a cDNA sequence (SEQ ID NO:9) aligned with the corresponding amino acid residues for human BAG-5 protein (SEQ ID NO:10).

Figure 6A shows the full length cDNA sequence for C. elegans BAG-1 protein (SEQ ID NO:11).

Figure 6B shows the 210 amino acid sequence for C. elegans BAG-1 protein (SEQ ID NO:12).

Figure 7A shows the full length cDNA sequence for C. elegans BAG-2 protein (SEQ ID NO:13).

Figure 7B shows the 458 amino acid sequence for $20\ C.\ elegans\ BAG-2$ protein (SEQ ID NO:14).

Figure 8A shows the full length cDNA sequence for $S.\ pombe\ BAG-1A$ protein (SEQ ID NO:15).

Figure 8B shows the 195 amino acid sequence for S. pombe BAG-1A protein (SEQ ID NO:16).

6

Figure 9A shows the full length cDNA sequence for S. pombe BAG-1B protein (SEQ ID NO:17).

Figure 9B shows the 206 amino acid sequence for $S.\ pombe\ BAG-1B$ protein (SEQ ID NO:18).

Figure 10 shows the topologies of the BAG-family 5 proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10); S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and C. elegans BAG-1 (SEQ ID 10 NO:12) and BAG-2 (SEQ ID NO:14). (A) The relative positions of the BAG domains are shown in black, ubiquitinlike regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like localization sequence are also shown. (B) The amino acid 15 sequences of the BAG domain for human BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10), S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18), and C. elegans BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) are aligned demonstrating 20 their homology. Black and gray shading represent identical and similar amino acids, respectively.

Figure 11 shows assays demonstrating the interaction of BAG-family proteins with Hsc70/ATPase. (A) Two-hybrid assays using yeast expressing the indicated 25 fusion proteins. Blue color indicates a positive interaction, resulting in activation of the *lacZ* reporter gene. (B) *In vitro* protein assays using GST-fusion proteins and ³⁵S-labeled *in vitro* translated proteins. (C) Co-immunoprecipitation assays using anti-Flag or IgG1 control antibodies and lysates from 293T cells expressing Flag-tagged BAG-1 (beginning at residue 116 of SEQ ID

NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), Daxx, or Apaf-1.

Figure 12 shows surface plasmon resonance analysis of BAG-family protein interactions SDS-PAGE analysis of purified Hsc70/ATPase. (A) recombinant proteins. (B) Representative SPR results of biosensor chips containing immobilized BAG proteins with and without maximally bound Hsc70/ATPase.

Figure 13 shows representative SPR results for 10 biosensor chips containing immobilized BAG-1 (beginning at residue 116 at SEQ ID NO:2), BAG-1(Δ C), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) proteins. Hsc70/ATPase was flowed over the chips (arrow/left) until maximal binding was reached (response units), then flow was continued without 15 Hsc70/ATPase (arrow/right). For BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6), Hsc70 was injected at 0.0175, 0.035, 0.07, 0.14, and $0.28 \mu M$.

Figure 14 shows BAG-family protein modulation of Hsc70 chaperone activity. (A) Protein refolding assay of 20 chemically-denatured luciferase by Hsc70 plus DnaJ in the absence or presence of BAG and BAG-mutant proteins. Concentration-dependent inhibition of Hsc70-mediated protein refolding by BAG-family proteins [BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEO ID NO:6)] but not by BAG-mutant (BAG-1 (Δ C). (C) Hsc70/Hsp40-mediated refolding of heat-denatured luciferase was assayed in the presence of (black bars) or absence of (striped bars) of 1.8 μ M Hip, with (lanes 3-10) or without (lanes 1,2) various BAG-family proteins (1.8µM) indicated (mean ±SE; n=3). A control (CNTL) is shown (lane 1) in which Hsc70 was replaced with an equivalent amount of BSA.

8

Figure 15A shows an expanded cDNA sequence for human BAG-3 protein (SEQ ID NO:19).

Figure 15B shows the corresponding amino acid residues for the human BAG-3 protein (SEQ ID NO:20) of 5 Figure 15A.

Figure 15C shows the expanded cDNA sequence (SEQ ID NO:19) aligned with the corresponding amino acid residues for human BAG-3 protein of Figure 15A (SEQ ID NO:20).

10 Figure 16A shows an expanded cDNA sequence for human BAG-4 protein (SEQ ID NO:21).

Figure 16B shows the corresponding amino acid residues for the human BAG-4 protein of Figure 16A (SEQ ID NO:22).

Figure 16C shows the expanded cDNA sequence (SEQ ID NO:21) aligned with the corresponding amino acid residues for human BAG-4 protein of Figure 16A (SEQ ID NO:22).

Figure 17A shows an expanded cDNA sequence for 20 human BAG-5 protein (SEQ ID NO:23).

Figure 17B shows the corresponding amino acid residues for the human BAG-5 protein of Figure 17A (SEQ ID NO:24).

Figure 17C shows the expanded cDNA sequence (SEQ 1D NO:23) aligned with the corresponding amino acid residues for human BAG-5 protein of Figure 17A (SEQ ID NO:24).

9

Figure 18 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), expanded BAG-3 (SEQ ID NO:20), expanded BAG-4 (SEQ ID NO:22), expanded BAG-5 (SEQ ID NO:24); 5 S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and C. elegans BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like nuclear localization sequence are also shown.

Definitions

The term "apoptosis", as used herein, refers to the process of programmed cell death, although not all programmed cell deaths occur through apoptosis, as used 15 herein, "apoptosis" and "programmed cell death" are used interchangeably.

The term "tumor cell proliferation", as used herein refers to the ability of tumor cells to grow and thus expand a tumor mass.

The term "cell migration", as used herein refers to the role cell motility plays in the invasion and potentially metastasis by tumor cells.

The term "metastasis", as used herein refers to the spread of a disease process from one part of the body to another, as in the appearance of neoplasms in parts of the body remote from the site of the primary tumor; results in dissemination of tumor cells by the lymphatics or blood vessels or by direct extension through serious cavitites or subarachnoid or other spaces.

10

The term "steroid hormone receptor function", as used herein refers to physiological, cellular and molecular functioning of receptors sites that bind with steroid hormones.

5 The term "substantially purified", as used herein, refers to nucleic acid or amino acid sequence that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Nucleic acid molecule" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single or double stranded, and represent the sense or antisense strand.

"Hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The terms "complementary" or "complementarity", 20 as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T binds to the complementary sequence "T-C-A".

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid and is referred to using the functional term "substantially homologous." The inhibition of

11

hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency.

The term "antisense", as used herein, refers to nucleotide sequences which are commplementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which 15 permits the synthesis of a complementary strand. introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further 20 transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is used in reference to the antisense, sometimes "positive" is sometimes used in reference to the sense strand.

"Amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited herein this term excludes an amino acid sequence of a naturally occurring protein. "Amino acid sequence", "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

12

The term "functional fragments" or "fragments", as used herein, with regard to a protein refers to portions of that protein that are capable of exhibiting or carrying out the activity exhibited by the protein as a whole. The portions may range in size from three amino acid residues to the entire amino acid sequence minus one amino acid. For example, a protein "comprising at least a functional fragment of the amino acid sequence of SEQ ID NO:1", encompasses the full-length of the protein of SEQ ID NO:1 and portions thereof.

A "derivative" of a BAG protein, as used herein, refers to an amino acid sequence that is alterd by one or more amino acids. The derivative may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., substitution of an apolar amino acid with another apolar amino acid (such as replacement of leucine with isoleucine). The derivative also have "nonconservative" changes, wherein a substituted amino acid has different but sufficiently similar structural or chemical properties that permits such a substitution without adversely effecting the desired biological activity, e.g., replacement of an amino acid with an uncharged polar R group with an amino acid with an apolar R group (such as replacement of glycine with tryptophan), or alternatively replacement of an amino acid with a charged R group with an amino acid with an uncharged Polar R group (such as replacement of lysine with asparagine).

10

15

13

Amino Acids - Apolar R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
alanine	methyl	ala	А
valine	2-propyl	aal	V
leucine	2-methylpropyl	leu	L
isoleucine	2-butyl	ile	I
proline	propyl* - cyclized	pro	P
phenylalanine	benzyl	phe	F
trytophan	3-indolylmethl	tyr	W
methionine	methylthioethyl	met	М

Amino Acids - Uncharged Polar R Groups

Amino Acid	Radical	Abbreviations		
		3-Letter	l-Letter	
glycine	Н	gly	G	
serine	hydroxymethyl	ser	S	
threonine	1-hydroxyethyl	thr	Т	
cysteine	thiolmethyl	cys	С	
tyrosine	4-hydroxyphenylmethyl	tyr	Y	
asparagine	aminocarbonylmethyl	asn	N	
glutamine	aminocarbonylethyl	gln	Q	

20 Amino Acids - Charged R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
aspartic acid	carboxymethyl	asp	D
glutamic acid	carboxyethyl	glu	E
lysine	4-aminobutyl	lys	K
arginine	3-guanylpropyl	arg	R
histidine	4-imidazoylmethyl	his	Н

25

5

10

14

Similar minor modifications may also include amino acids deletions or insertions or both. Guidance in determining which amino acid residues may be modified as indicated without abolishing the above desired biological 5 functionality may be determined using computer programs well known in the art, for example, DNASTAR software. addition, the derivative may also result from chemical modifications to the encoded polypeptide, including but not limited to the following, replacement of hydrogen by an 10 alkyl, acyl, or amino group; esterification of a carboxyl group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative. Further a derivative may also result from the substitution of a Lconfiguration amino acid with its corresponding Dconfiguration counterpart. 15

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of a protein/polypeptide or portions thereof (such as BAG-1) and, as such, is able to effect some or all of the actions of BAG-1 protein.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al., Anticancer Drug Des. 8:53-63 (1993)).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a family of BAG-1 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1S beginning at residue 116 of SEQ ID NO:2, BAG-2 (SEQ ID

15

NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEO ID NO: 8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate C.elegans [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast S.pombe NO:16), BAG-1B (SEQ ΙD [BAG-1A (SEQ ΙD acid full length amino sequences specifically the comprising human BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), and BAG-2 (SEQ ID NO:4) C. elegans BAG-1 (SEQ ID NO:12), and BAG-2 (SEQ ID NO:14), and S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); 10 and partial sequences comprising human BAG-3 (SEQ ID NO: 6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) and functional fragments thereof. In particular, the invention provides the amino acid sequences comprising human BAG-2 (SEQ ID 15 NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) proteins.

Another aspect of the present invention provides
the nucleic molecule and nucleotide sequences that encode
the family of BAG-1 related proteins from humans [BAG-1
(SEQ ID NO:1), BAG-2 (SEQ ID NO:3), BAG-3 (SEQ ID NO:5) and
(SEQ ID NO:19), BAG-4 (SEQ ID NO:7) and (SEQ ID NO:21) and
BAG-5 (SEQ ID NO:9) and (SEQ ID NO:23)], the invertebrate
C.elegans [BAG-1 (SEQ ID NO:11), BAG-2 (SEQ ID NO:13)] and
the fission yeast S.pombe [BAG-1A (SEQ ID NO:15), BAG-1B
(SEO ID NO:17)].

BAG-1L (SEQ ID NO:2) is a multifunctional protein that blocks apoptosis, promotes tumor cell metastasis, and contributes to factor-independent and p53-resistant cell growth. BAG-1L (SEQ ID NO:2) interacts with several types of proteins, including Bc1-2, some tyrosine kinase growth

16

factor receptors, steroid hormone receptors, and the p53-induced cell cycle regulator Siah-1A.

BAG-1 is a regulator of Hsc70/Hsp70 family molecular chaperones. A carboxyl-terminal domain in this 5 protein binds tightly to the ATPase domains of Hsc70 and Hsp70 ($K_n = 1$ nM) (Zeiner, M., Gebauer, M., and Gehring, U., EMBO J. **16**: 5483-5490, (1997)). BAG-1 modulates the activity of these molecular chaperones, acting as an apparent functional antagonist of the Hsp70/Hsc70associated protein Hip (3-5) (Höhfeld, J. and Jentsch, S., 10 EMBO J. 16: 6209-6216, (1997); Takayama, S., Bimston, D. N., Matsuzawa, S., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R. J., and Reed, J. C., EMBO J. 16: 4887-96, (1997); Zeiner, M., Gebauer, M., and Gehring, U., EMBO J. 16: 5483-5490, (1997)). In general, protein refolding is accomplished by Hsp70/Hsc70 through repeated cycles of target peptide binding and release, coupled to hydrolysis (Ellis, R., Curr Biol. 7: R531-R533, (1997)). BAG-1 appears to promote substrate release, whereas Hip stabilizes Hsp70/Hsc70 complex formation with target 20 peptides (Höhfeld, J., Minami, Y., and Hartl, F.-U., Cell. **83**: 589-598, (1995)). Since each substrate interaction with Hsc70/Hsp70 is unique in terms of the optimal length of time the protein target should remain complexed with Hsc70/Hsp70 for achieving new conformations, the net effect 25 of BAG-1 can be either enhancement or inhibition of the refolding reaction.

The 70kd heat shock family proteins (Hsp70/Hsc70) are essential to a variety of cellular processes and have been implicated in cancer, yet it is unclear how these proteins are regulated *in vivo*. A variety of co-chaperones have been identified which may target Hsp70/Hsc70 to different subcellular compartments or promote their

17

interactions with specific protein or protein complexes. BAG-1 appears to represent a novel Hsp70/Hsc70 regulator which differs functionally from all other mammalian cochaperones identified to date, such as members of the DnaJ-, Hip-, Hop-, and cyclophilin-families of proteins.

Another aspect of the present invention provides the amino acid sequence of a binding domain of about 40 to 55 amino acids that bind the a Hsc70/Hsp70 ATPase domain. The BAG domain is situated near the C-terminus, and the ubiquitin-like domains are situated near the N-terminus.

The BAG family of proteins of the present invention contain a common conserved C-terminal domain (the "BAG" domain) that facilitates binding to the ATPase domain of Hsp70/Hsc70. The carboxyl-terminal domain of BAG-1 binds to the ATPase domain of Hsc70/Hsp70 and regulates its chaperone function by acting as a ADP-ATP exchange factor. Other domains of BAG-1 mediate interactions with proteins such as Bcl-2 and retinoic acid receptors (RARs), allowing BAG-1 to target Hsc70/Hsp70 to other proteins, presumably modulating their function by changing their conformations.

Human BAG-1 was previously shown to inhibit Hsc70/Hsp70 dependent refolding of denatured protein substrates in vitro (S. Takayama, et al., EMBO J 16, 4887-96 (1997); M. Zeiner, M. Gebauer, U. Gehring, EMBO J. 16, 5483-5490 (1997); and J. Höhfeld, S. Jentsch, EMBO J. 16, 6209-6216 (1997)). In Example III, Part A the effects of recombinant human BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were compared using in vitro protein refolding assays similar to those employed previously for assessing BAG-1. The study showed that addition of equimolar amounts of each of the recombinant proteins to Hsc70 resulted in significant inhibition of luciferase refolding, with BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) showing somewhat

18

greater inhibitor activity than BAG-1 (Figure 4A). In a separate luciferase folding study BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) once again displayed inhibition of luciferase refolding, however in this study varying amounts of BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were added relative to Hsc70 which resulting in concentration-dependent inhibition of Hsc70 chaperone activity, i.e., luciferase folding (Example III Part A). Additional follow on studies using the same experimental protocols as the previous studies, as taught in Example IIA, have shown that BAG-4 (SEQ ID NO:22) also undergoes association with Hsc70/ATPase.

Yet another aspect of the present invention provides a nucleotide sequence having at least about 15 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides that can hybridize or is complementary under relatively stringent conditions to a portion of the nucleic acid sequences shown in Figures 1-9 and Figures 15-17, in particular the BAG domain as shown in in Figure 1B, e.g., nucleotides 552-593 of human BAG-3, or nucleotides 167-221 of human BAG-4.

Yet another aspect of the present invention 25 provides a compound of the formula,

$$R^{N} - R^{1}X^{1}R^{2}X^{2}R^{3}X^{3}R^{4}X^{4}R^{5}X^{5}R^{6}X^{6}R^{7}X^{7}X^{8}R^{9}X^{9}R^{10}X^{10}R^{11}X^{11} - R^{C}$$

wherein,

15

20

 R^N is a group of 1 to 552 independently selected amino acids;

 R^1 is a group of 3 independently selected amino acids:

 ${\rm X}^1$ is an amino acid with a charged or uncharged R group, such as aspartic acid, glutamic acid, asparagine, or glutamine;

 ${\ensuremath{\mathtt{R}}}^2$ is a group of 7 independently selected amino 5 acids:

 $\mbox{\ensuremath{X^2}}$ is an amino acid with a charged R group, such as glutamic acid;

 ${\ensuremath{R}^3}$ is a group of 5 independently selected amino acids;

10 X^3 is an amino acid with an apolar R group, such as leucine, methionine, or isoleucine;

 \mathbb{R}^4 is a group of 3 independently selected amino acids;

 ${
m X}^4$ is an amino acid with charged R group, such as aspartic acid or glutamine acid;

 R^5 is a single independently selected amino acid; X^5 is an amino acid with apolar or uncharged R group, such as leucine, valine, methionine, alanine or threonine;

20 \mathbb{R}^6 is a group of 15 independently selected amino acids;

 ${\tt X}^6$ is an amino acid with a charged or uncharged R group, such as arginine, lysine, glutamine or aspartic acid:

 R^7 is a group of 2 independently selected amino acids;

 χ^7 is an amino acid with a charged R group, such as arginine;

 X^8 is an amino acid with a charged R group, such as arginine or lysine;

 $\ensuremath{\,\text{R}^{9}}$ is a group of 2 independently selected amino acids;

 ${\tt X}^9$ is an amino acid with an apolar R group, such as valine;

R¹⁰ is a group of 3 independently selected amino acids;

20

 ${ exttt{X}}^{10}$ is an amino acid with an uncharged R group, such as glutamine;

 \mathbb{R}^{11} is a group of 2 independently selected amino acids;

 \mathbf{X}^{11} is an amino acid with an apolar R group, such as leucine; and

 $\ensuremath{\,\text{R}^{\text{C}}}$ is a group of 1 to 100 independently selected amino acids.

A nucleotide sequence of at least about 15 generally, about nucleotides 25 nucleotides. and, 10 preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides can be useful, for example, as a primer for the polymerase chain reaction (PCR) or other similar reaction mediated by a polymerase such as a DNA or RNA polymerase (see PCR 15 Protocols: A guide to methods and applications, ed. Innis et al. (Academic Press, Inc., 1990), which is incorporated herein by reference; see, for example, pages 40-41). addition, such a nucleotide sequence of the invention can be useful as a probe in a hybridization reaction such as 20 Southern or northern blot analysis or in a binding assay such as a gel shift assay.

A nucleotide sequence of the invention can be particularly useful as an antisense molecule, which can be DNA or RNA and can be targeted to all or a portion of the 5'-untranslated region or of the 5'-translated region of a bag-1 nucleic acid sequence in a cell. For example, an antisense molecule can be directed to at least a portion of the sequence shown as the BAG domain in Figure 1A, e.g., nucleotides 272-319 of human BAG-1L (SEQ ID NO:1), or nucleotides 79-147 of human BAG-5 (SEQ ID NO:9). Since the 5'-region of a nucleic acid contains elements involved in the control of expression of an encoded protein, an antisense molecule directed to the 5'-region of a nucleic

21

acid molecule can affect the levels of protein expressed in a cell.

A nucleotide sequence of the invention also can be useful as a probe to identify a genetic defect due a 5 mutation of a gene encoding a BAG protein in a cell. Such a genetic defect can lead to aberrant expression of a BAG protein in the cell or to expression of an aberrant BAG protein, which does not properly associate with a Bcl-2-related protein or Hsc70/Hsp70 protein in the cell. As a 10 result, a genetic defect in a gene encoding, for example, human BAG-1 can result in a pathology characterized by increased or decreased levels in protein folding.

Further a nucleotide compound or composition as taught in the present invention can be synthesized using routine methods or can be purchased from a commercial In addition, a population of such nucleotide source. sequences can be prepared by restriction endonuclease or mild DNAse digestion of a nucleic acid molecule that contains nucleotides as shown in the nucleotide sequences shown in Figures 1-9 and Figures 15-17 that encodes the amino acids sequences also shown in Figures 1-9 and Figures 15-17. Methods for preparing and using such nucleotide sequences, for example, as hybridization probes to screen a library for homologous nucleic acid molecules are well known in the art (see, for example, Sambrook et al., Molecular Cloning: A laboratory manual (Cold Spring Harbor Laboratory Press 1989); Ausubel et al., Current Protocols in Molecular Biology (Green Publ., NY each of which is incorporated herein by reference).

A particular nucleotide sequence can be designed based, for example, on a comparison of the nucleic acid molecules encoding any one of the BAG family proteins, as shown in Figures 1-9 and Figures 15-17, with another in the family. Such a comparison allows, for example, the

22

preparation of a nucleotide sequence that will hybridize to a conserved region present in both nucleic acid molecules, thus providing a means to identify homologous nucleic acid molecules present in other cell types or other organisms. In addition, such a comparison allows the preparation of a nucleotide sequence that will hybridize to a unique region of any of the BAG family nucleotide sequences, such as those corresponding to the BAG domain, thus allowing identification of other proteins sharing this motif. this regard, it is recognized that, while the human BAG-3 10 proteins shown as Figures 3 and 20, and human BAG-5 proteins shown as Figures 5 and 24, are only partial sequences, a variant human BAG-3 or BAG-5 produced, for example, by alternative splicing can exist and can be identified using an appropriately designed nucleotide 15 sequence of the invention as a probe. Such useful probes readily can be identified by inspection of the sequences shown in the disclosed Figures by a comparison of the encoding nucleotide sequences.

If desired, a nucleotide sequence of the invention can incorporate a detectable moiety such as a radiolabel, a fluorochrome, a ferromagnetic substance, a luminescent tag or a detectable binding agent such as biotin. These and other detectable moieties and methods of incorporating such moieties into a nucleotide sequence are well known in the art and are commercially available. A population of labelled nucleotide sequences can be prepared, for example, by nick translation of a nucleic acid molecule of the invention (Sambrook et al., supra, 1989; Ausubel et al., supra, 1989).

One skilled in the art would know that a method involving hybridization of a nucleotide sequence of the invention can require that hybridization be performed under relatively stringent conditions such that nonspecific background hybridization is minimized. Such hybridization

23

conditions can be determined empirically or can be estimated based, for example, on the relative GC content of a sequence and the number of mismatches, if known, between the probe and the target sequence (see, for example, Sambrook et al., supra, 1989).

invention further provides antibodies specific for human BAG family protein. As used herein, the includes polyclonal "antibody" and monoclonal antibodies, as well as polypeptide fragments of antibodies that retain a specific binding activity for human BAG-1 of 10 at least about 1 x 10^5 M⁻¹. One skilled in the art would know that anti-BAG-1 antibody fragments such as Fab, F(ab') and Fv fragments can retain specific binding activity for human BAG-1 (beginning at residue 116 of SEQ ID NO:2) and, 15 thus, are included within the definition of an antibody. In addition, the term "antibody" as used herein includes naturally occurring antibodies as well as non-naturally occurring antibodies and fragments that retain binding activity such as chimeric antibodies or 20 antibodies. Such non-naturally occurring antibodies can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains as described by Huse 25 et al., Science 246:1275-1281 (1989), which is incorporated herein by reference.

One skilled in the art would know that purified BAG family protein, which can be prepared from natural sources or synthesized chemically or produced recombinantly, or portions of a BAG family protein, including a portion of human BAG family protein such as a synthetic peptide as described above, can be used as an immunogen. Such peptides useful for raising an antibody include, for example, peptide portions of the N-terminal 85 amino acids or the BAG domain of any of the human BAG

24

proteins (see Figure 1B). A particularly advantageous use of such a protein is for the immunostaining, wherein the methods provides a process to contrast the immunostaining of BAG-family proteins in carcinoma cells with adjacent non-neoplastic prostatic epithelial and basal cells which are generally present in the same tissue sections. These results would be correlated with a Gleason grade to determine whether any of the BAG-family proteins tend to be expressed at higher or lower levels in histologically advanced tumors. From this process a determination can be made as to degree at which the disease is progressing in a given patient, i.e., a prognosis can be made.

Non-immunogenic fragments or synthetic peptides of BAG proteins can be made immunogenic by coupling the hapten to a carrier molecule such bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH), as described in Example IV, below. In addition, various other carrier molecules and methods for coupling a hapten to a carrier molecule are well known in the art and described, for example, by Harlow and Lane, Antibodies: A laboratory manual (Cold Spring Harbor Laboratory Press, 1988), which is incorporated herein by reference.

EXAMPLES

The following examples are given to enable those skilled in the art to more clearly understand and to practice the present invention. They should not be considered as limiting the scope of the invention, but merely as being illustrative and representative thereof.

15

25

EXAMPLE I

Isolation and Characterization of BAG-family cDNA Sequences

This example describes methods for isolating and characterizing of BAG-family cDNA sequences from human, nematode and yeast.

A. Cloning of human BAG cDNA sequences

Yeast two-hybrid library screening of a human Jurkat cell cDNA library was performed as described by 10 Takayama et al., <u>EMBO J.</u>, 16:4887-96 (1997); Matsuzawa et al., EMBO J., 17:2736-2747 (1998), which are incorporated herein by reference) using EGY48 strain yeast transformed with pGilda-Hsc70/ATPase (67-377 amino acids) and the lacZ reporter plasmid pSH18-34. Of the resulting ~5 x 106 15 transformants, 112 Leu colonies were obtained after 1 week incubation at 30°C. Assay of β -galactosidase (β -gal) activity of these colonies resulted in 96 clones. Mating tests were then performed using RFY206 yeast strain transformed with pGilda, pGilda mBAG-1 (1-219), or pGilda 20 Hsc70/ATPase. Of these, 66 displayed specific interactions with Hsc70/ATPase. The pJG4-5 cDNAs were recovered using KC8 E. coli strain which is auxotrophic for tryptophan DNA sequencing revealed 3 partially overlapping human BAG-1, 4 identical and one overlapping cDNAs encoding 25 BAG-2, and 2 partially overlapping BAG-3 clones.

Using the above described yeast two-hybrid screen with the ATPase domain of Hsc70 as "bait", several human cDNAs were cloned which encode portions of BAG-1 or of two other BAG-1-like proteins which are termed BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6). The longest of the cDNAs for BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) contained open reading frames (ORFs) of 207 and 162 amino acids, respectively, followed by stop codons. All BAG-1 (SEQ ID

26

NO:1), BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) cDNAs obtained by two-hybrid library screening with Hsc70/ATPase contained a conserved domain of about 40-50 amino acids which are termed the "BAG" domain and are shown in Figure 10. These results demonstrate that a family of BAG-1-related proteins all contain a conserved ~45 amino acid region near their C-terminus that binds Hsc70/Hsp70.

B. Identification of additional BAG-family proteins

the bBLAST and FASTA search programs also identified human ESTs that provided sequences for further investigation of BAG-family proteins. The putative BAG-4 (SEQ ID NO:8) and BAG-5 (SEQ ID NO:10) proteins contain BAG-domains that share the greatest sequence similarity with the BAG-domain of BAG-3 (SEQ ID NO:6). These were designated BAG-4 (Accession number AA693697, N74588) and BAG-5 (Accession number AA456862, N34101). BAG-4 has 62% identity and ~81% similarity to BAG-3, and BAG-5 has 51% identity and ~75% similarity to BAG-3.

Additional BAG-family orthologues or homologues 20 were also identified using computer-based searches and resulted in BAG-family homologue in the nematode C. elegans and the fission yeast S. pombe. The C. elegans genome encodes two apparent BAG-family proteins, which are most similar in their overall sequences to the human BAG-1 25 (Afo39713, gi:2773211) (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) (Afo68719, gi:3168927). The S. pombe contains two BAG-family proteins that share the greatest sequence similarity with human BAG-1 (Alo23S54,gi/3133105 30 and Alo23634, qi/3150250). The human and C. elegans BAG-1 proteins as well as S. pombe BAG-1A all have ubiquitin-like domains near their N-termini (see Figure 10A) of unknown function.

The overall predicted amino acid sequences of the C. elegans BAG-1 (SEQ ID NO:12) and S. pombe BAG-1A (SEQ ID NO:16) proteins are ~18% identical (~61% similar) and ~17% identical (~64% similar), respectively, to human BAG-1, implying origin from a common ancestral gene. elegans BAG-1 protein (SEQ ID NO:12), however, contains a 5 to 7 amino acid insert in its BAG-domain as compared to the human, murine, and yeast BAG-1 homologues (see Figure 10B), and is more similar to BAG-2 (SEQ ID NO:4) in regard to its BAG-domain. C. elegans and human BAG-2 also may be derived from a common ancestor as the C-terminal 225 amino acid region which encompasses both the BAG domain and upstream region of both C. elegans and human BAG-2 share ~34% amino acid sequence identity and ~70% similarity. human BAG-2 protein (SEQ ID NO:4), however, contains a 9 15 amino acid insert in its BAG-domain compared to C.elegans counterpart (see Figure 10B). Evolutionary-tree prediction algorithms suggest that human and C. elegans BAG-2 represent a distinct branch of the BAG-family that is more evolutionarily distant from the other BAG-family 2.0 None of the predicted BAG-family proteins proteins. contain recognizable regions analogous to those found in other Hsc70 regulatory proteins, such as the J-domains and family proteins G/F-domains of DnaJ Tetratricopeptide Repeat (TR) domains of Hip/Hop family proteins.

C. Yeast two-hybrid assay of BAG binding to Hsc70/ATPase

The longest of the cDNAs obtained for the BAG-2 and BAG-3 proteins were expressed with N-terminal transactivation (TA) domains in yeast and tested by yeast two-hybrid assay for interactions with fusion proteins consisting of Hsp70/ATPase or a variety of unrelated proteins (Fas, Siah, Fadd) containing N-terminal LexA DNA-binding domains. TA-BAG-2 and TA-BAG-3 demonstrated

28

positive interactions with LexA-Hsc70/ATPase, resulting in transactivation of a lacZ reporter gene that was under the control of LexA operators (Figure 11A). No interactions with LexA-Fas (cytosolic domain), LexA-Siah, LexA-Fadd, or LexA were detected (see Figure 11A) demonstrating that the BAG-2 and BAG-3 proteins interact specifically with Hsc70/ATPase. Specific two-hybrid interactions between Hsc70/ATPase and either BAG-2 or BAG-3 were also observed when BAG-2 and BAG-3 were expressed as LexA DNA-binding domain fusion proteins and Hsc70/ATPase was fused with a TA domain (see Figure 11A; right panel). These results demonstrate that similarly to BAG-1, BAG-2 and BAG-3 specifically interact with Hsc70/ATPase.

In order to determine whether the BAG proteins are capable of forming heterodimers, coexpression of BAG-2 and BAG-3 in the yeast two-hybrid assay was also performed. Coexpression of BAG-2 and BAG-3 failed to show interaction with BAG-1 or a deletion mutant of BAG-1 (\Delta C) which is missing part of its C-terminal domain required for Hsp70/Hsc70 binding suggest that these proteins do not form heterdimers.

D. Isolation and characterization of the complete open reading frame sequences of BAG-2 and BAG-3

In order to deduce the complete ORFs of BAG-2 and 25 BAG-3, a λ-phage cDNA library was screened as follows, using hybridization probes derived from the two-hybrid screening. A human jurkat T-cell λ-ZapII library cDNA library (Stratagene) was screened by hybridization using ³¹P-labeled purified insert DNA from the longest of the human BAG-2 (clone #11) and human BAG-3 (clone #28) cDNA clones. From about one million clones screened, 38 BAG-2 and 23 BAG-3 clones were identified, cloned, and their cDNA inserts recovered as pSKII plasmids using a helper phage method (Stratagene). DNA sequencing of λ-phage derived

29

human BAG-2 cDNA clones revealed an ORF encoding a predicted 211 amino acid protein, preceded by an in-frame stop codon. The longest human BAG-3 λ -phage cDNA clone contains a continuous ORF of 682 amino acids followed by a stop codon, but without an identifiable start codon (see Figure 10A).

Although BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) all contain a homologous BAG domain near their C-terminus, the N-terminal regions of these proteins are dissimilar. Using a combination of search tools (Prosite Search: PP search, using the Prosite pattern database, BCM Search Launcher, Baylor College of Medicine, and Blocks Search), it was determined that the BAG-2 N-terminal region contains potential kinase phosphorylation sites but otherwise shares no apparent similarity with other proteins or known functional domains.

In contrast, the predicted N-terminal region BAG-3 contains a WW domain as shown in Figure 10A. WW domains have been identified in a wide variety of signaling proteins, including a Yes kinase adaptor protein (YAP), the Na'-channel regulator Nedd4, formin-binding proteins, dystrophin, and the peptidyl prolyl cis-trans-isomerase Pin-1. These roughly 40 amino acid domains mediate protein interactions and bind the preferred peptide ligand sequence xPPxY (Sudol., TIBS, 21: 161-163, 1996, which is incorporated herein by reference).

30

EXAMPLE II

In vitro Association of BAG proteins and Hsc70/ATPase

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) bind Hsc70/ATPase in various in vitro assays.

A. Solution binding assay of BAG-2 and BAG-3 to Hsc70/ATPase

Association of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) with Hsc70/ATPase was determine by an in vitro 10 protein binding assay where Hsc70/ATPase or BAG-family proteins were expressed in bacteria as Glutathione S-Transferase (GST) fusion proteins. Purified cDNA sequences encoding residues 5 to 211 of human BAG-2 (clone #11) and the C-terminal 135 amino acids of human BAG-3 (clone #28) 15 (see Figure 10A) were subcloned into the EcoRI/Xho I sites of pGEX4T-1 prokaryotic expression plasmid (Pharmacia; Piscataway, NJ). These plasmids as well as pGEX4T-1-BAG-1, pGEX-4T-1-BAG-1 (Δ C), and pGEX-4T-1-XL which have been described previously (Takayama et al., supra (1997); Xie et 20 al., <u>Biochemistry</u>, 37:6410-6418, (1998), which are incorporated herein by reference), were expressed in XL-1 blue strain E. Coli (Stratagene, Inc., La Jolla, CA). Briefly, a single colony was inoculated into 1L of LB media 25 containing 50 $\mu g/ml$ ampicillin and grown at 37°C overnight. culture was then diluted by half with fresh LB/ampicillin and cooled to room temperature for 1 hr, before inducing with 0.4mM IPTG for 6 h at 25°C.

Cells were recovered and incubated with 0.5 mg/ml lysozyme in 50 mM Tris (pH 8.0), 150 mM NaCl, 1% Tween-20, 0.1% 2-mercaptoethanol, 5 mM EDTA, 1 mM PMSF and a mixture

31

of other protease inhibitors obtained from Boehringer Mannheim (1697498) at room temperature for 0.5 h, followed Cellular debris were pelleted by sonication. centrifugation at 27,500g for 10 min and the resulting 5 supernatants were incubated with 30 ml of glutathionine-Sepharose (Pharmacia) at 4° C overnight. The resin was then washed with 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, and 0.1% 2-mercaptoethanol until the OD 280nm reached <0.01. For removal of GST, the resin with immobilized GST-10 fusion protein was incubated with 10U of thrombin (Boehringer, Inc.) at 4°C in 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, 0.1% 2-Mercaptoethanol, and 2.5 mM CaCl2 overnight. Released proteins were then purified on Mono Q (HR10/10, Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0 and dialyzed into chaperone assay buffer.

The ability of BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) to bind Hsc70/ATPase in solution was then examined. GST control or GST-BAG proteins were immobilized on glutathione-Sepharose and tested for binding to 35Slabeled invitro translated (IVT) proteins. Immunoprecipitation and in vitro GST-protein binding assays were performed as described by Takayama et al., supra (1997), using pCI-Neo flag or pcDNA3-HA into which human 25 Bag-2 (clone #11) or human BAG-3 (clone #28) had been subcloned for in vitro translation of 35S-L-methionine labeled proteins or expression in 293T cells. As shown in Figure 11B, 35S-Hsc70/ATPase bound in vitro to GST-BAG-1, GST-BAG-2, and GST-BAG-3 but not to GST-BAG-1(Δ C) several other control proteins. BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) also exhibited little or no binding to themselves or to each other, demonstrating that these proteins do not strongly homo- or hetero-dimerize or 35 oligomerize. It should be noted, however, that BAG-2 (SEQ

32

ID NO:4) displayed weak interactions with itself in binding assays and produced a positive result in yeast two-hybrid experiments, demonstrating that it can have the ability to self-associate.

5 B. Binding of BAG proteins to Hsc70 in vivo

The ability of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins to interact in cells with Hsc70 was tested by expressing these proteins with N-terminal Flag epitope tags in 293T human epithelial cells using codescribed previously immunoprecipitation assays as 10 (Takayama et al., supra (1997)). cDNAs encoding the λ phage cloned regions of BAG-2 and BAG-3 were subcloned inframe into pcDNA3-Flag. Anti-Flag immune complexes prepared from 293T cells after transfection with plasmids 15 encoding Flag-BAG-1, Flag-BAG-2, or Flag-BAG-3 were analyzed by SDS-PAGE/immunoblot assay. As shown in Figure 10C, antiserum specific to Hsc70 detected the presence of BAG proteins associated with Hsc70, whereas control immunecomplexes prepared with IgG1 as well as anti-Flag immune 20 complexes prepared from cells transfected with Flag-tagged control proteins, Daxx and Apaf-1, did not contain Hsc70 associated protein. These results further demonstrate that BAG-family proteins specifically bind to Hsc70.

C. BIAcore assay of BAG protein binding to the ATPase 25 domain of Hsc70

BAG-1 (beginning at residue 116 of SEQ ID NO:2) is known to bind tightly to the ATPase domain of Hsc70 (Stuart et al., <u>J. Biol. Chem.</u>, In Press (1998)). BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins were therefore, examined for their ability to bind to Hsc70/ATPase. The affinity and binding kinetics of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) to Hsc70/ATPase was also compared to that of BAG-1 (beginning at residue 116 of

SEQ ID NO:2) for Hsc70/ATPase, using a surface plasmon resonance technique (BIAcore) which has been described previously (Stuart et al., *supra*, (1998) which is incorporated herein by reference).

BAG-family proteins were produced in bacteria and 5 purified to near homogeneity as shown in Figure 12A and in Example I. The purified BAG-1 described above (beginning at residue 116 of SEQ ID NO:2), -2 (SEQ ID NO:4), and -3 (SEQ ID NO:6) proteins were then immobilized on biosensor chips and tested for their interactions with 10 Hsc70 in the soluble phase. Kinetic measurements were performed using a BIAcore-II instrument with CM5 sensor chip and Amine Coupling Kit (Pharmacia Biosensor AB, Sweden). Briefly, for immobilization of proteins, the sensor chip was equilibrated with HK buffer (10 mM Hepes 15 (pH 7.4), 150 mM KCL) at 5μ l/min, then activated by injecting 17 μ l of 0.2M N-ethyl-N'-(3-diethylaminopropyl)carbodiimide and 0.05M N-hydroxysuccinimide followed by 35 μ l of the protein of interest, in 10 mM acetate, pH 3.5-4.5. Excess NHS-ester on the surface was deactivated with 17 μ l 1M ethanolamine-HCL (pH8.5). After immobilization, 5μ l of regeneration buffer (50 mM phosphate (pH 6.8) and 4M GuHCl) was injected. For binding assays, Hsp70 (Sigma, H8778) was dissolved in HK buffer, injected at 10 μ l/min across the prepared surface at various concentrations. The surface was regenerated after each injection with 5 μ l of regeneration buffer. The rate constants κ_{ass} and κ_{diss} were generated with BIAevaluation softward 3.01 (Pharmacia Biosensor AB). Addition of Hsc70 30 to chips containing BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) resulted in concentration-dependent binding, as reflected by an increase in the Response Units (RU) measured at the chip surface (shown in Figure 3B). In contrast, Hsc70 failed to display interactions in BIAcore assays with a 35 variety of control proteins as well as a mutant of BAG-1

34

lacking a C-terminal portion of the BAG domain which is required for Hsc70-binding (Figure 3B). Furthermore, flowing of various control proteins such as GST, BSA and Bcl-XL over the BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) chips resulted in negligible interaction. These results further demonstrate the specificity with which BAG-family proteins interact with and bind to Hsc70.

The rates of Hsc70 binding to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 10 (SEQ ID NO:6) proteins were similar, following pseudo first-order kinetics with estimated association rate constants (K_a) of 2.1, 2.1 and 2.4 x $10^5~\text{M}^{-1}~\text{sec}^{-1}$, respectively. After allowing binding of Hsc70 immobilized BAG-1 (beginning at residue 116 of SEQ ID 15 NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to reach plateau levels, the chaperone was removed from the flow solution and the dissociation rate was monitored. BAG-1 (beginning at residue 116 at SEQ ID NO:2) and BAG-2 (SEQ ID NO:4) exhibited similar dissociation rates, with relatively 20 slow loss of Hsc70 from the chip surface, resulting in estimated dissociation rate constants (κ_d) of 3.0 and 5.0 x 10⁻⁴ sec⁻¹, respectively (see Figure 3B). In contrast, Hsc70 dissociated more rapidly from biosensor chips containing BAG-3 (see Figure 3B), yielding an estimated κ_d of 1.7 x 10^{-3} 25 ${\tt sec}^{-1}$. From the kinetic data, the apparent affinities (${\tt K}_{\tt D}$ = K_d/K_d) were calculated for binding of Hsc70 to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) and were estimated to equal about $K_D=1.4 nM$, $K_D=2.4 nM$, and $K_D=7.4 nM$, respectively. These results demonstrate that the interactions of BAG-family with Hsc70 occur with apparent affinities proteins sufficient for physiological relevance.

35

EXAMPLE III

BAG-family proteins inhibit Hsp70/Hsc70-dependent protein folding

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins inhibit Hsp70/Hsc70-dependent refolding of denatured proteins similarly to a BAG-1 (beginning at residue 116 of SEQ ID NO:2) protein.

The effects of BAG-2 (SEQ ID NO:4) and BAG-3 (SEO ID NO:6) protein on Hsp70/Hsc70-dependent protein refolding was determined using in vitro protein refolding assays 10 similar to those described previously by Takayama et al., supra, 1998: Terada et al., <u>J Cell Biol.</u>, 139:1089-1095, 1997, which are incorporated herein by reference. Briefly, luciferase (20 μ M) was denatured in 25 mM Hepes-KOH, pH 7.2, potassium acetate, 5 mM DTT, 6M quanidine mM 15 hydrochloride at ~25°C for 1 h. Denatured luciferase was diluted 1:40 into 25 mM Hepes-KOH, pH 7.2, 50 mM potassium acetate, 5 mM DTT. Hsc70 (1.8 μ M), DnaJ (StressGen, Inc.) $(0.9\mu M)$, and various purified recombinant proteins as indicated were added to refolding buffer (30 mM Hepes-KOH, 20 pH 7.6, 120 mM potassium acetate, 3mM magnesium acetate, 2 mM DTT, 2.5 mM ATP) with 0.2 volume of diluted denatured luciferase to a final concentration of 0.1 μM . Luciferase activity was measured after 1.5 hr incubation at 35°C .

The combination of Hsc70 and DnaJ resulted in ATP-dependent refolding of chemically denatured firefly luciferase, with function of over half the denatured enzyme restored in a 90 minute reaction, as monitored by a chemiluminescence assay. In contrast, neither Hsc70 nor DnaJ alone were able to induce substantial refolding of denatured luciferase. Furthermore, little spontaneous

36

restoration of luciferase activity was observed with control proteins, BSA, GST or Bcl-XL (see Figure 4A).

Addition of recombinant purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to the above assays in amounts equimolar to Hsc70 (1.8 μ M) resulted in striking inhibition of luciferase refolding. BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) displayed somewhat greater inhibitory activity than BAG-1 (beginning at residue 116 of SEQ ID NO:2) as shown in Figure 4A. In contrast, the BAG-1 (Δ C) protein, which fails to bind Hsc70 as well as several other control proteins, had no effect on luciferase refolding.

In an additional refolding assay, described previously by Minami et al., J Biol. Chem. 271:19617-24, 15 1996), purified Hsc70 and human DnaJ homolog Hdj-1 (Hsp 40) used with additional cofactors provided in reticulocyte lysates (5% v:v) to produce a system capable of refolding denatured luciferase. Briefly, additional cofactors included, recombinant Luciferase (Promega: 20 QuantiLum TM), that had been heat denatured at 42°C for 10 min, 1.8 μM Hsc70 (Sigma; purified from bovine brain), 0.9 μM Hsp40, and various recombinant purified proteins. Luciferase activity was measured (Promega luciferase assay kit) using a luminometer (EG&G Berthold, MicroLumat 25 luminometer, Model #LB96P). All results were normalized relative to non-denatured luciferase that had been subjected to the same conditions. Control reactions lacking ATP, Hsc70, or Hsp40 resulted in negligible luciferase refolding. 30

Various amounts of purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6), relative to amounts of Hsc70 were used in the above-described protein refolding assay. Addition of BAG-family proteins resulted in a concentration-dependent

37

inhibition of Hsc70 chaperone activity. Furthermore, the BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) inhibition of Hsc70 chaperone activity was demonstrated to be as potent as that observed for BAG-1 (beginning at residue 116 of SEQ ID NO:2). In contrast, the BAG-1 (ΔC) mutant as well as other control proteins did not suppress Hsc70-mediated refolding of denatured luciferase. These results indicate that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) can inhibit Hsc70/Hsp70 dependent protein refolding activity to the same extent as BAG-1 (beginning at residue 116 of SEQ ID NO:2).

B. BAG competes with Hip for binding to Hsc70.

It is known that BAG-1 competes with Hip for binding to Hsc70, with these proteins exerting opposite effects on Hsc70-mediated protein refolding (Hohfeld, J., and Jentsch, S., Embo J., 16:6209-6216, 1997, which is incorporated herein by reference). In order to determine whether BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) also compete with Hip for binding to Hsc70, refolding assays were performed as described above in the presence of Hip protein.

Hip was purified as His6-protein. The fusion protein was induced from pET28-Hip (V. Prapapanich et al., Mol Cell Biol., 18:944-952, 1998, which is incorporated 25 herein by reference) with 0.1 mM IPTG at 25°C for 6h in BL21 cells. Cells from 1L of culture were resuspended into 50 ml of 50 mM Phosphate buffer (pH 6.8), 150 mM NaCl, and 1% (v/v) Tween-20 and then incubated with 0.5 mg/ml lysozyme 25°C for 0.5h, followed by sonication. After 30 centrifugation at 27,500g for 10 min, the resulting supernatant was mixed with 15 ml nickel resin (Qiagen, Inc.) at 4°C for 3 h with 25 mM imidazol. The resin was then washed with 50 mM phosphate buffer (pH 6.8), 25 mM imidazol, 150 mM NaCl and 0.1% Tween-20 until the OD280nm

38

reached a value of <0.01. His,-Hip protein was eluted with 250 mM imidazol in washing buffer (Qiagene, Inc.) and purified on Mono Q (HR10/10 Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0, followed by dialysis in chaperone assay buffer.

In the refolding assay reactions, addition of purified Hip at equimolar concentrations relative to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) (1.8 μ M) completely negated the inhibitory effects of the BAG-family proteins on refolding of denatured luciferase (see Figure 4C). These results demonstrate that the suppression of Hsc70 chaperone activity by BAG-family proteins is reversible, and that Hip antagonizes the effects of not only BAG-1 (beginning at residue 116 of SEQ ID NO:2), but also of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6).

In summary, these results demonstrate that BAG-family proteins all contain a conserved BAG domain near their C-terminus that binds Hsc70/Hsp70, and that human BAG-family proteins can bind with high affinity to the ATPase domain of Hsc70 and inhibit its chaperone activity through a Hip-repressable mechanism.

EXAMPLE IV

EXPANDED NUCLEIC ACID AND AMINO ACID SEQUENCES FOR HUMAN BAG-3, BAG-4 AND BAG-5

Following the procedures disclosed herein, the nucleic acid and amino acids sequences to human BAG-3, BAG-4 and BAG-5 were further expanded. The expanded sequences for BAG-3, BAG-4 and BAG-5 are shown in Figures 15, 16 and 17, respectively, with their respective sequence identification numbers, "SEQ ID NO"s.

10

15

39

We claim:

1. A compound of the formula,

 $R^{N} - R^{1}X^{1}R^{2}X^{2}R^{3}X^{3}R^{4}X^{4}R^{5}X^{5}R^{6}X^{6}R^{7}X^{7}X^{8}R^{9}X^{9}R^{10}X^{10}R^{11}X^{11} - R^{C}$

	wherein,	
5		R^N is a group of about 1 to 552 independently
		selected amino acids;
		R^1 is a group of 3 independently selected amino
		acids;
		X^1 is an amino acid with a charged or uncharged
10		R group;
		R^2 is a group of 7 independently selected amino
		acids;
		X ² is an amino acid with a charged R group;
		R^3 is a group of 5 independently selected amino
15		acids;
		X ³ is an amino acid with an apolar R group;
		R ⁴ is a group of 3 independently selected amino
		acids;
		X4 is an amino acid with charged R group;
20		R ⁵ is a single independently selected amino acid;
		X^5 is an amino acid with apolar or uncharged R
		group;
		R ⁶ is a group of 15 independently selected amino
		acids;
25		X^{b} is an amino acid with a charged or uncharged
		R group;
		R ⁷ is a group of 2 independently selected amino
		acids;
		X is an amino acid with a charged R group;
30		X^8 is an amino acid with a charged R group;
		R ⁹ is a group of 2 independently selected amino
		acids;

 X^9 is an amino acid with an apolar R group;

40

R¹⁰ is a group of 3 independently selected amino acids;

X¹⁰ is an amino acid with an uncharged R group;
R¹¹ is a group of 2 independently selected amino acids;

 X^{11} is an amino acid with an apolar R group; and R^{C} is a group of about 1 to 100 independently selected amino acids.

- 2. A substantially purified nucleic acid molecule having a nucleotide sequence corresponding to or complementary to at least 20 nucleotides from a nucleotide sequence selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:21) and (SEQ ID NO:23).
- 3. The nucleic acid of claim 2 having a nucleotide sequence corresponding to or complementary to a nucleotide sequence that encodes a functionally active BAG family protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:22) and (SEQ ID NO:24).
- 4. The nucleic acid of claim 3 selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).
- 5. The nucleic acid of claim 3 complementary to a nucleotide sequence that encodes a functionally active BAG protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).
 - 6. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:3).

- 7. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:5).
- 8. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:7).
- 9. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:9).
 - $$10.\ A$$ substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:19).
- 10 11. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:21).
 - $$12.\ A$$ substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:23).
- 13. A substantially purified BAG family protein encoded by the nucleic acid molecule of claim 1.
- 14. A substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24) or a fragment, a derivative or a mimetic thereof.
 - 15. A substantially purified protein corresponding to the amino acid sequence of 157 to 204 of (SEO ID NO:2).
- 25 16. A substantially purified protein corresponding to the amino acid sequence of 272 to 319 of (SEQ ID NO:2).

42

- 17. A substantially purified protein corresponding to the amino acid sequence of 164 to 211 of (SEQ ID NO:4).
- 18. A substantially purified protein 5 corresponding to the amino acid sequence of 418 to 510 of (SEC ID NO:20).
 - 19. A substantially purified protein corresponding to the amino acid sequence of 378 to 457 of (SEQ ID NO:22).
- 10 20. A substantially purified protein corresponding to the amino acid sequence of 6 to 97 of (SEQ ID NO:24).
- 21. A substantially purified protein corresponding to the amino acid sequence of 180 to 257 of 15 (SEQ ID NO:24).
 - 22. A substantially purified protein corresponding to the amino acid sequence of 272 to 349 of (SEQ ID NO:24).
- 23. A substantially purified protein 20 corresponding to the amino acid sequence of 362 to 444 of (SEQ ID NO:24).
 - 24. A pharmaceutical composition comprising a nucleic acid molecule of claim 1 useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.
 - 25. A method of modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function by administering a nucleic acid molecule of claim 1.

- 26. A pharmaceutical composition comprising a substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24), or a fragment, a derivative or a mimetic thereof, useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.
- 10 27. A method of modulating tumor cell proliferation by administering a pharmaceutical composition of claim 26.
- 28. A method of modulating cell migration and metastasis by administering a pharmaceutical composition of claim 26.
 - 29. A method of modulating steroid hormone receptor function by administering a pharmaceutical composition of claim 26.
- 30. A substantially purified antibody that 20 specifically binds to a BAG family protein of claim 14.
 - 31. The antibody of claim 30, wherein said antibody is a monoclonal antibody.

44

32. A method for detecting the presence of a BAG family protein in a sample, comprising the steps of:

- a. obtaining the sample;
- b. adding to said an antibody of claim 11 under suitable conditions for the binding of said antibody with the BAG family protein; and
- c. detecting said bound BAG family protein.
- 33. A method for detecting the presence of a first nucleic acid molecule that encodes a BAG family protein in a sample, comprising the steps of:
 - a. obtaining the sample;
 - b. adding to said sample a second nucleic acid molecule capable of hybridizing with said first nucleic acid molecule under suitable conditions for the binding of said second nucleic acid molecule with said first nucleic acid molecule; and
 - c. detecting said hybridized first and second nucleic acid molecules.
- 34. A method of determining the risk of metastatic spread of cancer or prognosis of cancer patients by determining the level of expression of a BAG-family protein.

5

15

ACCOCCCC CACCTICCAT	CACTICCAL	cacrossess	TCAACAATTC	CCCCCCCCC TCAG	CCCCCC R G	COCCCCCA CARR	CACCOCCACOS PRG	CCACCT:3CAG D R E	06
concrement conscrints R L G S R L R	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	COCCUTICOS A L R	CCACACACAC	Actorica P R Q	GTCCHCCCC S E P	CCCCCACC	G P P	recentras P S R	180
CONCOCCIO COCCOGENIAC RPPARST	CCCCCACIAC R S T	TOCCACOCOS A S G	CATCACCAC H D R P	CCACCACCOCC TRG	COCCOCCOCC A A A A	OCCOCTOCCA G A R R	COCCOCCAL PRM	GAAGAAGAAA K K K BAG-1M	270
Accessor octosaces TRRRSTR	OCTOCACCOS S T R	CACCCACCAG	TICHOCOCH L T R S	COCACCACIT E E L	CACCCTCAGT T L S	GACCAACCCA E E A T	CCTCCAGTCA W S E	ACACCICACC E A T	360
CACATACAC ACCOCACCA Q S E E A T Q	ACCCCACCCA A T Q	COCCAMONG C E E	ATCHATCCCA M N R S BAG-1	OCCACCACOT Q E V	CACCCCCCAC T R D	CACCACTOCA E E S T	CCCCCAACCCA R S E	GSAGGTGACC E V T	450
ACCEPCEARA TOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOCOC	TOOCOOCACC A A A A	Tasserchee G L T	∢ ⊢	CCCACACCAA H S N	TEACAAOCAC E K H	CACCTICATE D L H V	TIACCICCCA T S Q	0040010400 Q G S	270
ACTORACCOG TIGICCOACA S E P V V Q D	TIGICCAACA V Q D	ccroscopo L A Q	OTTOTTCAAC V V E E	AGGICATAGG V I G	OCTTCCACAG V P Q	TCTTTTCAGA S F Q K	AACTCATATT L I F	TAACCCAAAA K G K	630
TCTCTCAAGG AAATGCAAAC S L K E M E T		ACCONTOTICA P L S	OCACTICCAA A L G I	TACAACATOG Q D G	TTOCCOOSTC C R V	atgitaatig m l i g	ocaaaaaaa k k n	CAGTCCACAG S ? Q	720
CAACACOTTO AACTAAACAA E E V E L K K	AACTBAAGAA L K K	GTTCAAACAT L K H	TICCACAAGT L E K S	CTGTGGAGAA V E K	CATACCTCAC I A D	CACCTCCAAG Q L E E	acticaataa L n k	AGAGCTTBACT E L T	810
GRATICAC ASSITTICT	ACCONTINCT G F L	OCCCAACCAT	TICCACCIC L Q A E	AAGCTCTCTG A L C	CAAACTICAT K L D	ACCACACTAA R R V K	AACCCACAAT A T I	ACACCASTITE E Q F	06
ATCHACATOT TOCHOCACATO M K I L E E I	TOCHOCHCHI E E I	TEACACACTG D T L	ATCCTCCCAG I L P E	AAAATTTCAA N F K	ACACAGTAGA D S R	TTCAAAACCA L K R K	AAGGCITIGGI G L V	AAAAAAGGIT K K V	066
CACCOUNTIC TRACCIONATIC Q A F L A E C	TRACCORPAGIO A E C	TEACACAGIC D T V	COCCOMPCA E Q N I	TCTCCCACCA C Q E	CACTCACCCC T E R	CTOCAGTOTA L Q S T	CAAACTTTOC N F A	CCTCCCCC#G	1080
TCACCTCTAC CACAAAAACC	CHCHARARGO	crerecrece	CTGAAGAATG	COCCACCAG	creroceare	TCTCCATCCC	AATITIACCTG	АТТСТСОВ	1170
COCTOCTOCS	OCTOCTOS OCCACTOSC	CATTTCCCAA	TITICCIACT	CTCACACTOG	TTCTCAATGA	AAAATAGTGT	CTITICICATE	TGAG DAAAGC	1260
TCCTACTICTG	TECTATIONS TITTICAGA	AAAAAAAAA	æ						1291

FIGURE 2A

					1 1001	12.		
06	180	270	360	450	540	930	720	8 10
	СССТ ОС АБСС GAAOATCAAC K I N	OCTOOROCTC L E L	AAATAGCCAG N S Q	т 6 явто тся Е U S	TCTOOATGAT L D D	GTTTCAATCC F Q S	т васан весс р К А	ATCARGCTAT TRGROCATIC TARROGAGCT GOTTCCARAR CTCTGCARCA AARTGCTGAR AGCAGATTCA ATTAGTCTTC ARRCCTARGA 1 K L L E H S K G A O S K T L Q Q N A E S R F N
	616AC66C6A 166C1CA66C A Q A	GCCTGGACCA L D Q	ACAGTATCCA S 1 Q	CTCTCACCGT L T U	TCAATAAGTT N K F	TTGATCAGAA D Q K	TTGARARCTC E N S	ATTAGTCTTC
	с вссвсвтт в вяввсттявя н	CTGCTGGAGA L L E S	GAAATGATCC E M I H	ATGGGAAGAA H G R T	GATOAGGTGG D E U U	САТGGGCCAG И 0 Р U	CTTAGAAATA L A N 1	AGCAGATTCA S R F N
	6646666666 CCCC6C6TC6	CTCCAGCCGC S S R	ARTCCTTCTG	AAACCGTTTG N R L	AAGGATTATT R 1 1	TGAGGTGCCA E U P	AGAGACTCTG E T L	AAATGCTGAA N A E
	оссотсясяс стсттосстя	TGGCTGACCG A D R	ААСАСАААСА Е К Е	ATCTGACTGC L T A	AGCATGCCAC H A T	CATGITCAIC C S S	AGAGAAGATT R R L	CTCTGCAACA L Q Q
	TTGCCCCCGC GGCCGGTGAC	TCCTCCTCCA S S S M	OCTOTTORGC A U E Q	GAAGAATTAA E E L N	GARTCCCTAR E S L K	CTCTACABTG L Y S A	ABGARARTTA K K I K	GOTTCCARAB G S K T
	GCAGCCGCGG TGTCGCGAAG TCCTCCCGGG TTGCCCCCGC GGCGTCAGAG GGAGGGCGGG CGCCGCGTTG GTGACGGCGA CCCTGCAGCC CAAGGAGCGC TCCACTCGCT GCCGCCGGAG GGCCGGTGAC CTCTTGGCTA CCCCGCGTCG GAGGCTTAGA TGGCTCAGGC GAAGATCAAC	CTTCTGCCGC F C R	посттовно сттововно воспосностоттовос внововняем вытесттето онавтовтсе венетнестне внатросснова ${\sf R}$ в ${\sf R}$	GACRIGAGGC AGATCRGIGA CGGAGAAAGA GARGAATTAA ATCTGACTGC AAACCGTTTG ATGGGAAGAA CTCTCACCGT TGARGTGTCA D m r q i s d g e r e e l n l t a n r l m g r t l t u e u s	GTRORNACAR TTRORARCCC CCRGCRGCRR GRRTCCCTRA RGCRTGCCRC ARGGRTTATT GRTGRGGTGG TCRATRAGTT TCTGGRTGAT ${\sf U}$ E T I R N P Q Q Q E S L K H R T R I I D E U U N K F L D D	TTGGGRAFIG CCARGAGTCA TTTARTGTCG CTCTACAGTG CATGTTCATC TGAGGTGCCA CATGGGCCAG TTGATCAGAR GTTTCAATCC LONARY REST LANGERS ${\sf L}$ ${\sf Y}$ ${\sf S}$ ${\sf A}$ ${\sf C}$ ${\sf S}$ ${\sf S}$ ${\sf E}$ ${\sf V}$ ${\sf P}$ ${\sf H}$ ${\sf G}$ ${\sf P}$ ${\sf U}$ ${\sf D}$ ${\sf Q}$ ${\sf K}$ ${\sf F}$ ${\sf Q}$ ${\sf S}$	ATROTARITG GCTGTGCTCT TGARGATCAG AAGAAAATTA AGAGAGATT AGAGACTCTG CTTAGAAATA TTGAAAACTC TGACAAGGCC	TARAGGAGCT K G A
	тотсосовно тссястсост	ACGAGGGGGG E G R	CTTTGAGAGA L A E	AGATCAGTGA	TTAGAAACCC R N P	CCARGAGTCA K S H	GCTGTGCTCT C A L	TAGAGCATTC E H S
	6СРВССССССВ ТОТССССВНЯ ТССТСССВВО ТТВСССССВС ВВСВТСЯВВ ВВЯВЕСВВ СВССВСВТТВ ВТВЯСТВЕ ВТВЕСВВЕСВ СССТВСЯВСС СРЯВВЕТСЯВ ВВЕСТТЯВА ТВССТСЯВВС ВЯВВЯТСЯВС СРЯВВЯТСЯВС ТСРАВВЕТСВСТ В ССВЕСВВЕСВВЕСТ В ССВЕТИВВЕТЕ В В ССВЕТИВВЕТЕ В В В ССВЕТИВВЕТЕ В В В В В В В В В В В В В В В В В	GCTARRGCCA ACGRGGGGG CTTCTGCCGC TCCTCCTCCA TGGCTGARCCG CTCCRGCCGC CTGCTGGRGA GCCTGGACCA GCTGGAGCTC A K A N E G R F C R S S S M A D R S S R L L E S L D Q L E L	RGGGTTGRAG R V E A	GACATGAGGC D M R Q	OTAGARACAR U E T I	TTGGGAAATG L 0 N A	RTRGTRATTG	RTCRRGCTRT

3 / 35

FIGURE 2B

TCTTTGTTAG GTATAACCAC TTAGTTGACA		1080		9211
ССАТТТЯСЯС ЯЯТЯСЯСЯВО ОТОТЯЯЯВЯТ ОЯТЯЯЯЯТС ТЯТТТТЯВТТ ОЯТЯЯСТЯСТ ТСТТТОТТЯС СТЯТЯВССЯС ТТЯОТТОЯ СЯ	CTGRIRGITG TITCRGRIGA GGRARATRIT CCRICARGIA ICTICAGIII IGIGARIARC ARARCIAGCA AIRITITARI IRICIRICIA	GRGATITIT AGATIGARTI CITGICITGI ACTAGGAICI AGCAIAITIC ACTAIICIGI GGAIGARIAC ATAGILIGIG GGGAAAACAA	ACGITCAGCI AGGGGCAAAA AGCAIGACIG CIIIIICCIG ICIGGCAIGG AAICACGCAG ICACCIIGGG CAIIIAGIII ACIAGAAAII	CTTACTGG

1 I GOKE 3	
GOOGACTOC GORTOCANCE COGGOCGGG GOCANCTTCT CTGGACTGGR CC	AGRANGETTE CTAGCCGGCC AGTTGCTACC TCCCTTTATC 90
RELRIQP RAA AKFS GLO Q	KFLAGQLLPPFI
TECTECTTES CUTCIGOCAG COAGGAGGST ATTICCAGAG ACTICCAGOS CIT	CTCTGGCC ACGTCACCCC CCCCTTTAAT TCATAAAGGT 180
	LATSPPPLIHKG
GCCCGCCCC GCCTTCCCGG ACACGTCGGC GCCCGAGAGG GGCCCACGGC GG	
RRRRLPG HUG GGEG PTA A	AR PETR RPEPAP
CICACCICCI COCCACCICGI CAGACCCCAA CCCACCATGA CCCCCCCAC CC	ACTEGECE ATGATGEAGG TGGCGTCCGG CAACGGTGAC 360
	SP H H Q U A S G H G D
COCCHCCCTT TOCCCCCCGG RTGGGGGGATC RAGRITCGGCC CCCGGGCCCG CTG	GCCCTTC TTCGTGGACC ACAACAGCCG CACCACTACG 450
ROPL PPG HEIKIDP Q T G H	PFFUDH NSRTTT
TOGARCORCE DECOCOTOCC CTCTGRGGGC CCCRRGGRGA CTCCRTCCTC TO	CONATGGC CONTECCOGG AGGGCTCTAG COTGCCCCCT 540
UNDPRUPSEGPKET PSS R	NG PSREGSR L PP
GCTROGGRAG GCCACCCTGT GTRCCCCCAG CTCCGACCAG GCTACATTCC CAT	TTCCTGTG CTCCRTGAGG GCGCTGAGGA CCGGCAGGTG 630
AREG HPU YPQ LRPG YIP 1	PULHEGAENRQU
- CACCCTTTCC ATGTCTATCC CCAGCCTGGG ATGCAGCGAT TCCGAACTGA GGG	DESCRICER GEOGETECTS REAGGTECOR STERESTOTS 720
RPFH UYP QPG HORF RTE A	AA AAPO RSO SPL
	•
COCCCCRTCC CAGARACCAC TCACCCAGAT AAACAGTGTG CACAGGTGGC ACC	COGCOCCO COMPONDO COCCOMPONTO COMPONDACO 8 10
ROHPETT QPD KQCG QUA A	nn nn ur rh S h G r
GAGCGGTCCC AGTCTCCAGC TGCCTCTGAC TGCTCATCCT CATCCTCCTC GGC	CONSCRETE COTTOCTOCS GORGGAGCIAG COTGGGCAGT 900
ERSO SPA ASD CSSS SSS A	S L P S S G R S S L G S
CACCAGCTCC CGCGGGGTA CATCTCCATT CCGGTGATAC ACGAGCAGAA CGT	TRACCORG CORGORGOCC RECOCCITECTT CORGRANGOC 900
HOLPRGYISIPUINEONU	
n d c t ii o t i o t i e d ii o	

CAGAAGACCC ACTACCCAGC GCAGAGGGGT GAGTACCAGA CCCACCAGCC TGT	
QKTH YPA QRG EYQT HQP U	Y H K I Q G D D H E P R
COCCTGCCGG CGGCATCCCC GTTCAGGTCA TCTGTCCAGG GTGCATCGAG CCC	GGGGGGC TCACCAGCCA GGGGCGGCAC GCCACTCCAC 1170
PLRA A SP FRS SUOG ASS R	E G S P A R S S T P L H
TOCOCCTOSC CORTOCGTGT GCACACCGTG GTCGACAGGC CTCAGCAGCC CAT	IGACCCAT CGAGAAACTG CACCTGTTTC CCAGCCTGAA 1260
SPSPIRV HTU UDRP QQP H	THRETAPUS QPE
AACAAACCAG AAAGTAAGCC AGGCCCAGTT GGACCAGAAC TCCCTCCTGG ACF	ACATCOCA ATTOMAGTGA TOOGCAAAGA GGTGGATTCT 1350
NKPE SKP GPU GPEL PPG H	IP I Q U I R K E U D S
CONTROLL CONSCION CONTROL TOTAL AND TAXABLE CONTROL CO	
	TECCCCCT GCTCCAGTTC CTTGTCCTCC TCCCAGCCCT 1440
KPUSQKPPPPSEKUEUKU	TECCCCCT GCTCCAGTTC CTTGTCCTCC TCCCAGCCCT 1440

G P S R	TIGTOCCCTC TTCCCCCAAG	AGTGTGGCTA CAGAAGAGAG S U A T E E R	GCCACCCCC ACCACTOCCC A A P S T A P	CTGCAGAAGC TACACCTCCA 1530 A E A T P P
AAACCAGGAG A	PAGCCOAGGC TCCCCCAAAAA A E A P P K	CATCCAGGAG TOCTGAAAGT H P G V L K V	GGAAGCCATC CTGGAGAAGG E A I L E K V	TGCAGGGCT GGAGCAGGCT 1620 Q G L. E Q A
GTAGACAACT T	TTGAAGGCAA GAAGACTGAC EGKKTD	HAAAAGTACC TGATGATCGA KKYL HIE	ACAGTATTTG ACCAAAGAGC E Y L T K E L	TGCTGGCCCT GGATTCAGTG 1710 LALDS U
GACCCCGACG C	GACGAGCCGA TGTGCGTCAG R A D V R Q	GCCAGGAGAG ACGGTGTCAG A R R D G V R	CAAGGTTCAG ACCATCTTGG K U Q T I L E	AAAAACTTOA ACAGAAAGCC 1800 K L E Q K A
ATTGATGTCC C	CAGGTCAAGT CCAGGTCTAT	GAACTCCAGC CCAGCAACCT E L Q P S N L	TCAACCAGAT CAGCCACTCC E A 0 Q P L Q	AGCCARTCAT COACATOGGT 1890 A I H E H G
GCCGTGGCAG (CAGACAAGGG CAAGAAAAAT O K G K K H	CCTCCAMATG CACAMCATCC A G N A E D P	CONCACACAAA ACCCAGCAGC H T E T Q Q P	CAGANGCONC AGCAGCAGCG 1980 EATAAAA
ACTTONANCE O	CONGORGOROT GRICAGROPOC S S H T D T	CCTGGTANCC CAGCACCACC P G N P A A P	GTACCCTCTG CCCTGTAAAAA	GTCAGACTCG GAACCGATGT 2070
GTOCTTTAGG (CATTTAGTT COATCOATTT	CAGAGACTTT AGGTCAGTTG	GTTTTGATTA GCTGCTTGGT	
PROPOSICE ALTER A	ANGGOCTRAA AGOGANAATG	RECETTED TOPATATION	TACTOTTGTA CAATTANHOA	
CTTTRACCCC (вттесттетт стеслессст	сточноттов воноссоно	CACCIGITAG CTGTGGTTGT	
OOACTOOAGG (COTACATOGO CACATOAATTA	REPORTED TRANSPORTED PO	CATTIATORO ANATOTTOCC	
PORTETTER	ARTTAMARTA OCTOACTTTA	A DADAGAGTAA AATGTGCCAG	ONGOCORTHOU PHATATCTOTA	TOTTOGREGA CTTTRATOCT 2520 *2520

5 / 35

АСОЯТ ЯТССТ СТАВСЯССАЯ СВЯТТОСЯВО ОССЯВВОТТТ СВЯТТСТТЯТ АСАЯЯТССЯС ССТЯТООТСС АЯСЯТЯСССС ССЯБОСССТО
<u> вовсемянтяс тесстсятяс тсявовостт яттятосясс товттятяст сявяссявтт ястссясявя явттссяявт ясттяссотт</u>
сятствесяя сявсесяяет сеявтетете вттювятетя тесесявеяв вяствтеянь ястоянвеяе есестеттяя вводенватт
ссявоятятс свесттелея внясестввя вталесетве сесяттятес ттятволеят ввтятсвтя втелення в насявесв в $^{ m H}$
ACTOTACGAC CACAAGAAAG ATGCGTGGGC TTCTCCTGGT GCTTATGGAA TGGGTGGCCG
RCCACCCGGC ARTCICIACA TGACTGARG TACTICACCA TGGCCTAGCA GIGGCICICC A P C N L Y M I E S I S P W P S S G S P
GCCCARGGAT TCTTCATACC CCTATAGCCA ATCAGATCAA AGCATGAACC GGCACAACTT P Y S Q S D Q S M N B H N F
GGGGRCAGTG AACAATGATG ATTCAGATCT TTTGGATTCC CAAGTCCAGT ATAGTGCTGA GCCTCAGCTG TATGGTAATG CCACCAGTGA G T U N N D D S D L L D S Q U Q Y S A E P Q L Y G N A T S D
CCRTCCCRAC ARTCARGATC ARAGTAGCRG TCTTCCTGAR GRATGTGTAC CTTCAGATGA RAGTACTCCT CCGAGTATTA ARARARATCAT h p n n q d q s s s l p e e c u p s d e s t p p s i k k i i
АСЯТОТСТЕ САВАНЕСТСС АСТАТСТТОЯ АСАВОЯВТА ВАВОВЯТТЕ ТАВОВАВАНА В СЕТРОВИТИВ ТОТОВИВНИВ В СЕТОВОВИВИВИТИТЕ В R
ARTICCTARCC ARGGARCTIT TGGARCTGGA TICAGITIGAA ACTGGGGGCC AGGACTCTGT ACGGCAGCC AGARARAGAGG CTGTTTGTAA $_{ m H}$ L T $_{ m K}$ E L L $_{ m E}$ L L $_{ m E}$ L D $_{ m S}$ U E T G G Q D $_{ m S}$ U $_{ m R}$ Q A $_{ m R}$ K E A U C K

6	180	270	360	450	540	630	689
ATT008CR0T	ACATATATTG T Y I D	GCARACTTGT G N L S	AAGCAGCTGC K Q L L	ATTCTCAGCT	GCTTCTATGT	TCTGCTGCTT	
GROGABITARA ARRIGARCTI CICCAROCRC ARRACCARRACCA ARRACROPATI GCRGGGITTA ATTOGRCAGI $oldsymbol{E}$ I $oldsymbol{K}$ I $oldsymbol{E}$ I $oldsymbol{K}$ I $oldsymbol{E}$ I $oldsymbol{G}$ I $oldsymbol{G}$ I $oldsymbol{G}$ I $oldsymbol{G}$ I $oldsymbol{G}$ A $oldsymbol{E}$ I $oldsymbol{G}$ I $oldsymbol{G}$ A $oldsymbol{G}$ I $oldsymbol{G}$ A $oldsymbol{G}$ I $oldsymbol{G}$ A $oldsymbol{G}$ I $oldsymbol{G}$ A $oldsymbol$	TGGATGAGGT AAGTNTTGAA AAAAACCCCT GCATCCGGGA AGCCAGGAGA AGAGCAGTGA TCGAGGTGCA AACTCTGATC ACATATATTG DE USXEKNPCIS LITYID	ACTTGRAGGA GGCCCTTGAG ARABGARAGC TGTTTGCTTG TGAGGRGCAC CCATCCCATA ARGCCGTCTG GARCGTCCTT GGARACTTGT ${\sf L}$ K ${\sf L}$ K ${\sf R}$ K ${\sf L}$ F A ${\sf C}$ E E H P S H K A V M N V ${\sf L}$ G N ${\sf L}$ S	CTGAGATCCA GGGRGARGTT CTITCATITG ATGGARATCG ARCCGATARG ARCTACATCC GGCTGGARGA GCTGCTCRCC ARGCAGCTGC	ТАGCCCTGGR TGCTGTTGRT CCGCAGGGAG RRGAGAGTG ТАRGGCTGCC AGGRARCARG CTGTGRGGCT TGCGCRGRAT RTTCTCRGCT A L D A U D P Q G E E K C K R A R K Q A U R L A Q N I L S Y	ятстсвяест вявитствят вяятвевяет яствяянне сявявятете ясттттвятя ствтттвея сттеятятет всттетятвт	свотстсяют ятттятоятт сявосяяятт стяттсяютя тстостостт	
AAACAGAATT T E L	ТС6А66Т6СА Е V Q	янвссатств в и и	GGCTGGAAGA L E E	CTGTGAGGCT U R L	CTGTTTTGCA	GRAGCARATT	
CTGAGCTCCA L S S K	AGAGCAGTGA R A V I	CCATCCCATA P S H K	ARCTACATCC N Y I R	АССААРСАВС В К Q A	ясттттевтя	ATTTATGATT	ятсяяняня
TGAATTGTAC E L Y	AGCCAGGAGA A R R	тонооносно Е Е Н	AACCGATAAG T D K	тановствсс К А А	сявявятстс	сявтстсявт	TTCCATTCGG
AAAACCCTTC N P S	GCATCCGGGA I R E	TGTTTGCTTG F A C	ATGGAAATCG 0 N R	ARGAGAAGTG E K C	астоннятно	E Y . RIAC GIGCATAITI	ACGITARCTI
CTCCAAGCAC L Q A Q	AAAAACCCCT K N P C	AAAAGAAAGC K R K L	CTTTCATTTG L S F D	CCGCAGGGAG P Q G E	GARTGGGAGT	E W E Y TGATTTATAC	TCATTACAGE
ARATGARCTT N E L	AAGTNTTGAA S X E	ICTTGRAGGA GGCCCTTGAG RARAGI	GAGATCCA GGGAGAGGTT CTTTCA E 1 Q G E V L S	TGCTGTTGAT A U D	банатстват	L D L K S D E W ATRORGAGCT TTCAGTTCAT TGATTT	TIGRIGITGC ARGACABATA ICATTACAGC ACGITARCII TICCATICGG AICAAAAAA
OHORARARA E I K	TGGATGAGGT D E U	ACTTGAAGGA L K E	CTGAGATCCA E I Q	TAGCCCTGGA A L D	ятстсвясст	L D L ATAGAGAGCT	TTGATGTTGC

7 / 35

FIGURE 6A

WO 00/14106

PCT/US99/21053

8 / 35 FIGURE 6B

MKVNVSCSSV	QTTIDILEEN	QGEDESILTL	GQLRDRIATD	NDVDVETMKL	50
LHRGKFLQGA	DDVSLSTLNF	KENDKIIVMG	GKNALVDDAG	FKMLMQYEKH	100
NLSNLQKAYD	LNLRDVADLE	RGFLEKPKQV	EMGKKLEKKV	KYFNEEAERH	150
LETLDGMNII	TETTPENQAK	RNREKRKTLV	NGIQTLLNQN	DALLRRLQEY	200
QSVLNGDIPE					210

FIGURE 7A

ATGCCAGTCG	TGAACATACC	AATCAAAATA	CTTGGTCAGA	ATCAATCACA	50
TAGTCGAAGT	AACTCCTCGT	CTTCTGTTGA	CAACGATCGA	AATCAACCAC	100
CACAGCAGCC	ACCTCAACCG	CAACCACAAC	AGCAATCTCA	GCAACAATAC	150
			CATCATTCCA		200
			GGACTTTCCC		250
CTGGGTTCCC	AAACGATTCT	GAATGGTCTT	CGAATTTCCC	GTCGTTTCCA	300
AATTTCCCAA	GTGGATTCTC	AAATGGAAGT	TCTAATTTCC	CTGATTTTCC	350
AAGATTCGGA	AGAGATGGAG	GACTATCGCC	AAACCCACCG	ATGCAAGGAT	40 0
			AATCTCCAAC		450
			CCACAATATT		50 0
			TCCTCCGTCA		550
CATCTCGACC	ACCATCTCGT	ACTCGTGAAC	CAAAGGAACC	TGAGGTACCC	60 0
			GAGAAGAAGG		650
			GGGTGATGAG		70 0
			ATAATTGCGA		75 0
			CTAACCTCCC		800
			ACGTAATCAA		850
			ATAAAATTGA		900
			GAAATGGAAA		950
			TAACTGCATG		1000
			TCACTGACCG		1050
			ACTCCACGAA		1100
			GATCGATGAA		1150
			TGTGCCTACA		1200
			GCCACCTGCC		1250
			TCAGAAACGC		1300
			ATGCTGAGAG	AACGAAAGCA	1350
GATTTGATGG	ATGATCAAAG	CGAATAG			1377

10 / 35

FIGURE 7B

MPVVNIPIKI	LGQNQSHSRS	NSSSSVDNDR	NQPPQQPPQP	QPQQQSQQQY	50
	HHSNGFSPNF				100
NFPSGFSNGS	SNFPDFPRFG	RDGGLSPNPP	MQGYRRSPTP	TSTQSPTSTL	150
RRNSQQNQAP	PQYSQQQPQQ	AQQRQTTPPS	TKASSRPPSR	TREPKEPEVP	200
ERPAVIPLPY	EKKEKPLEKK	GSRDSGKGDE	NLEENIAKIT	IGKNNCELCP	25 0
EQETDGDPSP	LTSPITEGKP	KRGKKLQRNQ	SVVDFNAKTI	VTLDKIELQV	300
EQLRKKAAEL	EMEKEQILRS	LGEISVHNCM	FKLEECDREE	.IEAITDRLTK	350
RTKTVQVVVE	TPRNEEQKKA	LEDATLMIDE	VGEMMHSNIE	KAKLCLQTYM	400
NACSYEETAG	ATCQNFLKII	IQCAADDQKR	IKRRLENLMS	QIENAERTKA	450
DLMDDQSE					458

11 / 35

FIGURE 8A

ATGTCAGAAA AGACT	AGCAC AGTTACAAT	A CACTATGGAA	ATCAGCGATT	50
TCCGGTAGCA GTCAA	TCTAA ATGAGACGT	r AAGTGAACTG	ATTGATGATT	100
TACTIGAAAC GACTG	AGATT TCTGAGAAG	A AAGTCAAGCT	TTTTTACGCT	150
GGCAAGCGTT TAAAA	GACAA AAAAGCCTO	G TTATCAAAAT	TGGGTTTAAA	200
AAATCATAGT AAAAT				250
CCAAGGAAAA AGACA	CGGTT GAGCCCGCT	CGAAAGCGGA	AGCGGAGAAT	300
CCTGTATTTT CGCGT	CATTTC TGGAGAAAT.	A AAAGCCATCG	ATCAGTATGT	350
TGACAAAGAA CTTTC	CCCCA TGTACGACA	A TTACGTAAAT	AAACCGTCGA	400
ACGATCCAAA GCAGA	AAAAC AAACAGAAA	TAATGATAAG	TGAACTACTT	450
TTACAACAGC TTTTA	AAATT GGATGGAGT	r gacgtactgg	GCAGCGAGAA	50 0
ATTGCGTTTT GAACG	GAAGC AACTTGTTT	C TAAGATCCAA	AAAATGTTGG	550
ATCACGTTGA CCAAA	CAAGC CAAGAAGTG	G CCGCATAG		588

WO 00/14106

PCT/US99/21053

12 / 35

FIGURE 8B

MSEKTSTVTI	HYGNORFPVA	VNLNETLSEL	IDDLLETTEI	SEKKVKLFYA	50
GKRLKDKKAS					100
PVFSRISGEI	KAIDQYVDKE	LSPMYDNYVN	KPSNDPKQKN	KQKLMISELL	150
LQQLLKLDGV	DVLGSEKLRF	ERKQLVSKIQ	KMLDHVDQTS	QEVAA	195

FIGURE 9A

איזיין איזיי	TTACCCAGTT	GTGTTCTATG	GATAAAAAAT	ATTGGATCTC	50
TCTAGCTGTA	TTGTCAGTTA	CTGTTTTGAT	TAGCGCATTA	TTGAAAAAGA	100
GAGGTACTGA	AACCGAAGAT	ATTGTCGTTG	TTCATTACGA	TGGCGAAAAG	150
THEFT A ASTER	TGTTGCGACA	ACCAAGGCTG	AATATGGTTT	CTTACACTAG	200
THE COURT COLOR	CGCGTGTGCA	ACGCATTTTC	AGTAATGCCC	GACAAAGCGT	250
CTCTCAAGTT	AAACGGGGTG	ACCCTCAAGG	ATGGTTCACT	TTCCGACCAA	300
AAACSTESTEAA	ATGGAAGTGA	ATTAGAGCTC	GAATTACCCA	AACTGAGCCC	350
GGCAATGCAA	CAAATTGAAG	CATATATAGA	TGAGCTTCAA	CAGGATCTCG	400
TCCCTAAAAT	TGAAGCCTTC	TGCCAATCGT	CTCCCGCTTC	GGCACAAGAT	450
GTTCAAGATT	TGCATACACG	CCTTAGTGAA	ACATTGTTGG	CTAGGATGAT	500
AAAATTAGAT	GCTGTTAATG	TTGAAGACGA	CCCAGAAGCT	CGTCTTAAAA	550
CANAGAAGC	TATTCGTTTA	TCTCAACAAT	ATTTGAGTAA	ACTAGATTCC	600
	AAAACAAATG				621

14 / 35

FIGURE 9B

MSFFTQLCSM D	KKYWISLAV	LSVTVLISAL	LKKRATETED	IVVVHYDGEK	50
LNFVLRQPRL N	MVSYTSFLR	RVCNAFSVMP	DKASLKLNGV	TLKDGSLSDQ	100
NVQNGSELEL E	LPKLSPAMQ	QIEAYIDELQ	QDLVPKIEAF	CQSSPASAQD	150
VQDLHTRLSE T	'LLARMIKLD	AVNVEDDPEA	RLKRKEAIRL	SQQYLSKLDS	200
TKNQNK					206

FIGURE 10A

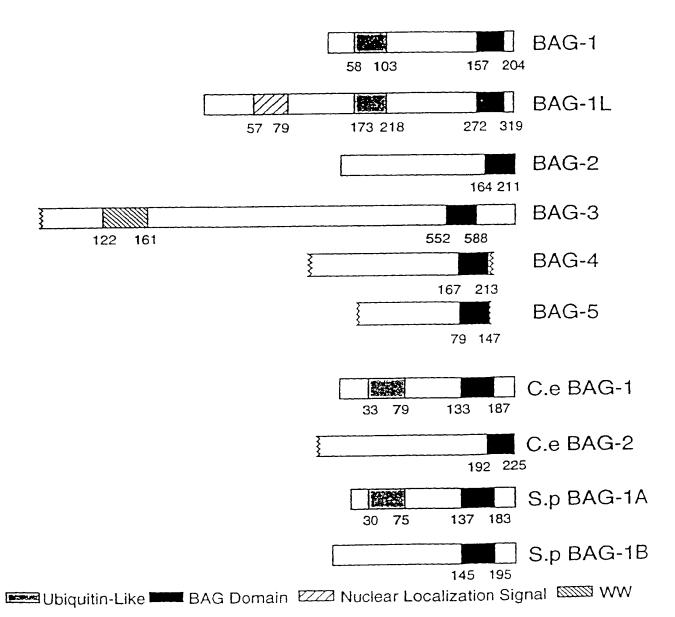
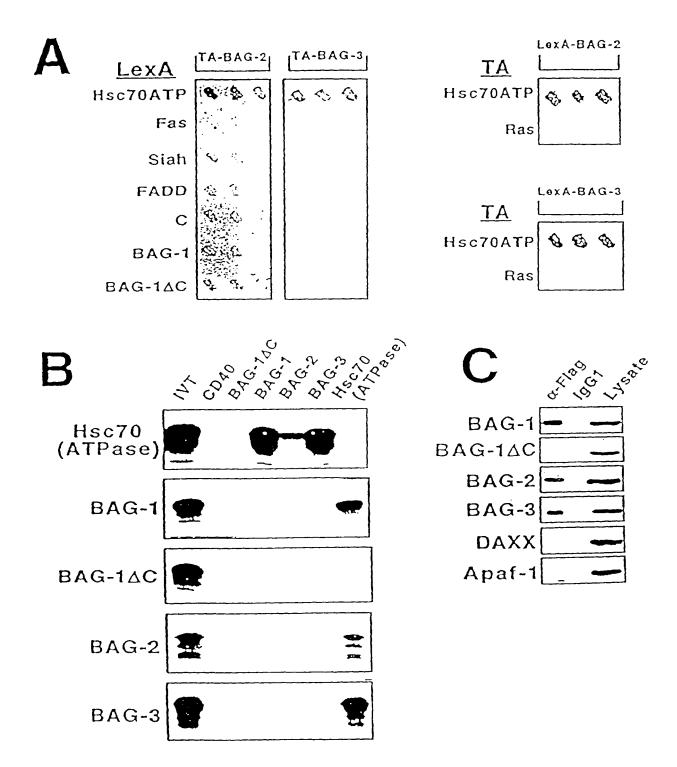


FIGURE 10B

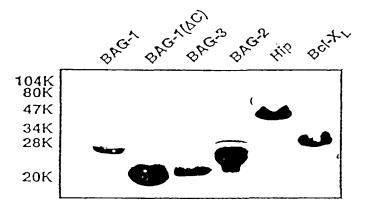
heag-1 heag-3 heag-4 heag-5 meag-1 C.e eag-1 S.p eag-13 S.p eag-13 C.e eag-2

11121141

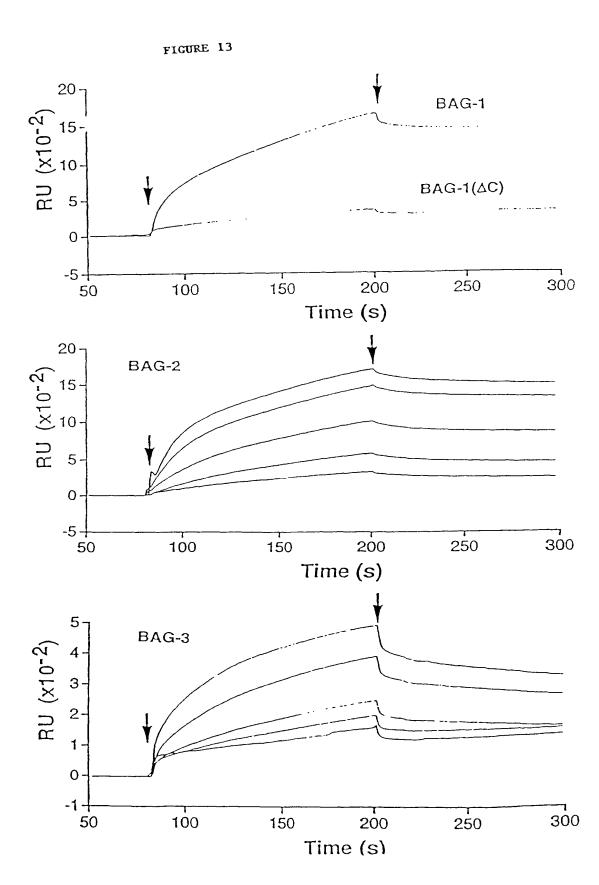
记器包含型品包括型



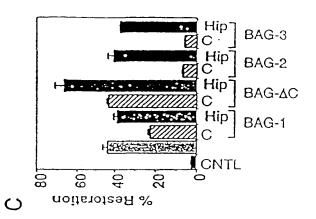
18 / 35 FIGURE 12

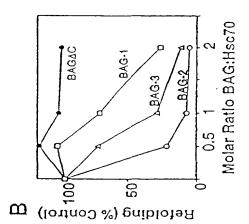


DEIGDOOID JEIO COLLEGER L.



20 / 35 FIGURE 14





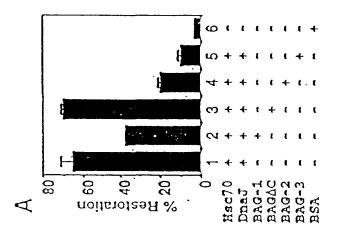


FIGURE 15A

GCGGAGCTCC GCATCCAACC CCGGGCCGCG GCCAACTTCT CTGGACTGGA	50
CCAGAAGTTT CTAGCCGGCC AGTTGCTACC TCCCTTTATC TCCTCCTTCC	100
CCTCTGGCAG CGAGGAGGCT ATTTCCAGAC ACTTCCACCC CTCTCTGGCC	150
ACGTCACCCC CGCCTTTAAT TCATAAAGGT GCCCGGCGCC GGCTTCCCGG	200
ACACGTCGGC GGCGCAGG GGCCCACGGC GGCGGCCCGG CCAGAGACTC	250
GGOGOCOGGA GOCAGOGOCO CGCACOCGCG COCCAGCGGG CAGACCCCAA	300
COCAGCATGA GOGOCGOCAC CCACTCGCCC ATGATGCAGG TGGCGTCCGG	350
CAACGGTGAC CGCGACCCTT TGCCCCCGG ATGGGAGATC AAGATCGACC	400
CGCAGACCGG CTGGCCCTTC TTCGTGGACC ACAACAGCCG CACCACTACG	450
TGGAAGGACC GGGGGGTGCC CTCTGAGGGC CCCAAGGAGA CTCCATCCTC	500
TGCCAATGGC CCTTCCCGGG AGGGCTCTAG GCTGCCGCCT GCTAGGGAAG	550
GCCACCCTGT GTACCCCCAG CTCCGACCAG GCTACATTCC CATTCCTGTG	009
CTCCATGAAG GCGCTGAGAA CCGGCAGGTG CACCCTTTCC ATGTCTATCC	650
CCAGCCTGGG ATGCAGCGAT TOCGAACTGA GGCGGCAGCA GCGGCTCCTC	200
AGAGGTCCCA GTCACCTCTG CGGGGCATGC CAGAAACCAC TCAGCCAGAT	750
AAACAGTGTG GACAGGTGGC AGCGGCGGCG GCAGCCCAGC CCCCAGCCTC	800
OCACGGACCT GAGCGGTCCC AGTCTCCAGC TGCCTCTGAC TGCTCATCCT	850
CATOCTOCTO GGOCAGOCTG CCTTCCTCCG GCAGGAGCAG CCTGGGCAGT	006
CACCAGCTCC OGOGGGGGTA CATCTCCATT CCGGTGATAC ACGAGCAGAA	950
CGTTACCCGG CCAGCAGCCC AGCCCTCCTT CCACAAAGCC CAGAAGACGC	1000
ACTACCCAGC GCAGAGGGGT GAGTACCAGA CCCACCAGCC TGTGTACCAC	1050
AAGATOCAGG GGGATGACTG GGAGOCOCGG CCCCTGCGGG CGGCATCCCC	1100
GTTCAGGTCA TCTGTCCAGG GTGCATCGAG CCGGGAGGGC TCACCAGCCA	1150
GGAGCAGCAC GCCACTCCAC TCCCCTCGC CCATCCGTGT GCACACCGTG	1200
GTCGACAGGC CTCAGCAGCC CATGACCCAT CGAGAAACTG CACCTGTTTC	1250
CCAGCCTGAA AACAAACCAG AAAGTAAGCC AGGCCCAGTT GGACCAGAAC	1300
TCCCTCCTGG ACACATCCCA ATTCAAGTGA TCCGCAAAGA GGTGGATTCT	1350

FIGURE 15A

20

MSAATHSPMM QVASGNGDRD PLPPGWEIKI DPQTGWPFFV DHNSRTTTWN

FIGURE 15B

DPRVPSEGPK ETPSSANGPS REGSRLPPAR EGHPVYPQLR PGYIPIPVLH 100
EGAENRQVHP FHVYPQPGMQ RFRTEAAAAA PQRSQSPLRG MPETTQPDKQ 150
CGQVAAAAAA QPPASHGPER SQSPAASDCS SSSSSASLPS SGRSSLGSHQ 200
LPRGYISIPV IHEQNVTRPA AQPSFHKAQK THYPAQRGEY QTHQPVYHKI 250
QGDDWEPRPL RAASPFRSSV QGASSREGSP ARSSTPLHSP SPIRVHTVVD 300
RPQQPMTHRE TAPVSQPENK PESKPGPVGP ELPPGHIPIQ VIRKEVDSKP 350
VSQKPPPPSE KVEVKVPPAP VPCPPPSPGP SAVPSSPKSV ATEERAAPST 400
APAEATPPKP GEAEAPPKHP GVLKVEAILE KVQGLEQAVD NFEGKKTDKK 450
YLMIEEYLTK ELLALDSVDP EGRADVRQAR RDGVRKVQTI LEKLEQKAID 500
VPGQVQVYEL QPSNLEADQP LQAIMEMGAV AADKGKKNAG NAEDPHTETQ 550
QPEATAAATS NPSSMTDTPG NPAAP

FIGURE 15C

EQUIDANCE CONTOCRACE CONSCIONAL ECONACTICE CINCACTURA CONCARCTE CINCONFOC ACTIVITIES TOCCITITES	9 0
TOUTCOTTOE COTUTOCON CONCENSEST ATTICONANC NOTTOCHOOC CTCTUTOCOC NOTTONOCOC COCCTITINAT TONTHANGET	100
ECCUPACION ECCUTATORIA MONOCITARES ECCENACION ECCUCACION ECCUCACIO	270
CECACCECCE COTCACCECE CACACCOOM COCACONTEA COCOCOCONC COACTOCOCE ATENTORICE TECCETOCCE CAACCETEAC	340
езнось вы поско	
COCCACCCTT TOCCCCCCC ATGGGGGGTC AGACTCCACC COCACCGCC CTGGCCTTC TTCTTGGGC ACAGCGCC CACGCCTGCC A B F L F F G W F I F W F E W F F F T T	450
TECHNOLOGY COCCOCTOCY CTCTURESOC COCAMOGRICA CTCCATCCTC TECHNATURE CCTTCCCCCC RECOCTCTRE ECTECCCCCCT	£ 4 0
ecthogonac eccaccete ethococcae eccheatic catectete etcoateage ecceptagna ecceptagna coecaccete a x I e κ 3 v κ 1 f κ 1 k 1 e κ 1 k 1 e κ 1 k 1 e κ 2 v κ	430
CACCOCTITICS ATSTETRICS CONSCISSES ATSCARGEAT TOCKNACTUR SECESCASES SOCIETY REMOTESTED A TRACTIC ACACCTUTE	720
CONCOUNTS CAGAMACCAS TOROCOMENT AMMOSTSTS CACACCTUCS ACCORDING CONCOUNTS CON	610
ENGOCETOCK NOTITIONACE TOCTTOTACE TOCTTOTTCT CATOCTTCTC COCONCOCTE OCTTOCTCCC COMMISSION OCTGECCNOT	+04
CHOCAGCITIC COCCCCCTR CHICTOMIT COCCTUMING MCCAGGNCAA CETTHCOCCC CONCONCCCC MCCOCTCCTT CCAGNAACCC M Q I Z I M K A	***
CACAMORACIC ACTRICICACI COMCRICICE CACTRICACA COCACCACCO TOTOTROCAC AMORTOXACE CICATRACTO COACCACCO Q K T K T T T A Q X C I T Q T K Q T V I K K I Q C) D V I T X	1040
COCCTUCERC COCCATODOC STRONGSTON TOTATOCAGE STECRATORIC COCCAGECC TONORACCA CORRECACIÓN COCACTOCAC L L N N J I I X J J V Q C N J J X I C J J N X J J T J L R	1170
TECCCTCCC CCATCCCTCT CONCACCCTC STCCACCCCC CTCACCACCT CCACCACACTC CACCACCTC CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACACT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACCT CACCACACACT CACCACACACT CACCACACACT CACCACACACT CACCACACACT CACCACACACA	1240
BACABACCAC BAGGTRACOC BOCCODACTI SCACCACAAC TOCOTOCTOC RCACATOCCA STITUBACTCA TOCOCAAACA SCITGATICT M K Y I J K Y \otimes 2 V \otimes 3 L L 2 2 \otimes 8 L Y I Q V I X K I V \otimes 3 J	1350
WARCTICHT COCROMOCC COCROCIOC ICTROMICS INCRESTOR WELLOCOCCE CLICAGIC CLICIOCIOC ICCOCCCCA	1440
CCCCTTCTC CTCTCCCCTC TTCCCCCAAC ACTCTCCCTX CACAACACAC CCCACCCCCC ACCACTCCCC CTCCACAACC TACACCTCCA	1510
WANCONCINC WICCORDOCC ACCORDANA DALCONCONC LOCLEWANGL CONFICUNCA C LOCHCOCCC ACCORDICAL CONFICUNCA CARRAGE LOCHCOCCA CARRAGE LOC	1420
ETHICACACCT TYCHNOCCAN GARACTURE AMARACTHOC TORTURTORS ACMICISTITIC RODAMCACC TOCTOCOCCT SCRITTCRCTC Y B K C	1710
ERCOCCURACE ERCEROCCUR TOTOCTTONE ECONCERCRE ROCETOTONE ERROCTTONE ROCETOTONE RAMACTTOR ROMERANGOE B F I C R A B V R Q A R R B C V R K V Q T F L I K L I Q K A	1400
ALLOWED SECTIONS CONCENTRAL ENVIRONCE CONCORDED LONG OF A LH L H ε	[490
ECCETCOOK CACACARCOC CHACAAAAAT ECTCOARTG CHCAACATOC CONCACHCAA ACCOACCACC CHCAACCORC ACCACCACCACACCACACACACACACACACACACACA	1440
ACTICANAC CONCACAT ENCACACIOS OCTOCOMOS CONCACACIÓN ESTACOCATO COCTODIAM ATORCACIÓN ENACOCATOS T 3 N 2 3 N 3 3 N 4 3 4 5 4 6 7 4 4 8 7 4 7 4 4 4 4 4 4 4 4 4 4	2070
CTCCTTRICC CANTITIBAC TECRITICAL ETCACACACT STRACTICACT SCCTTTIRS SECTOCITE CTRECOCCER ACTICCCTCC	2140
ACCCAMAGE CONTINUAC COCOMINAC CAMPICATE CTITICTET REPLICITING TOTORICAM TANGCACTI CCTTCTTTT	2250
PERCENTIONS TOURCOCCUTE ENTERCORES CANCENCES ACTIVED CONTROLL SETTEMPORT ENTERCASE SETTEMPORT STATEMPORT SETTEMPORT SETTE	23+0 2+00
WILLIAM ICCUMMIT WANTHOLD WILDRESS CHEMONAL LOCONCASC CHECCARE ELECTRICIT INTOWAST	2620
ANTOCIDICAT TITC	2514

FIGURE 16A

50 150 220 220 300 300 350 400 450 550 600 600 650 700 750 850 950 1000 1100 1150	
CGGTGGGAGC GGGGCGGAA GCGCTTCAGG GCAGCGGATC CCATGTCGGC CCTGAGGCGC TCGGGCTACG GCCCCAGTG CGGTCGCTC TACGGCCGCT ACTACGGGCC TCGGGGTGCA GATGTGCCG TACACCCACC TCCACCCTTA TATCCTCTTC GCCCTGAACC TCCCCAGGCCT CCCATTTCCT GGCGGTGCG CGGGGGCCCCCGGAACC TCCCCAGGCCT CCCATTTCCT GGCGGTGCG ATGGCTACTA TCCTCGGGA GCCGCTGGC CAGAGCCTGG TCGAGCCGCA GGAAGCCACC AGGAGCAGC ACCATATCCT ACCTACATT CTACTTATC GCAAGTCTACT GCGAGATCTA GGCCTCCTTA CCCAAGTACA TACCTGTAA GGACCCACAAT GCAGGCCAG AGTTTCAATC TTATCACAATT CTATCTCAGG GGTCCAACAT ACCCCCAGG CCCTGGGGCA AATACTGCCT CATACTCAGG GGTCCAACAT ACCCCCAGG CCCTGGGGCA AATACTGCCT CATACTCAGG GGTCCAACAT ACCCCCAGGT TAACTCAGAC CAGTTACTCC ACGAAGTTC CAAGTACTTA GCACCTGGTT ATACTCAGAC CAGTTACTCC ACGAAGTTC CAAGTACTTA GCACAGGCCTT CACGACCCCCT TTAGGGGGCA ATCTATCCC AGCAGGACTG TCAGAGTGTTC CACAATCAGG ACCGCCCATT ATCCTTATGG AGATGGTAAT CGTAGTGTTC CACAATCAGG ACCGCCCATT ACCTTATGG AGATGGTAAT CGTAGTGTTC CACAATCAGG ACCGCCCATT ACCTTATGG AGATGGTAAT CGTAGTGTTC CACAATCAGG ACCGCCACTCA CCCTTATCC TGGCCTTCAT CGTAGTGTTTC CACAATCAGG ACTGCCCAGTC ACATGACTACA AAGATGCTTA GCACCCCACGTCA CCCCTTCAC CCCCGTTCATCA GCAGCCCCACGTCA CCCCTTCACAATCAGA ACCGGCCCAAA CTTTCCTTGCATCAACAACACAAC	CTGGAGAGG TCCAGTATCT TGAACAAGAA GTAGAAAT TTGTAGGAAA AAAGACAGAC AAAGCATACT GGCTTCTGGA AGAAATGCTA ACCAAGGAAC

FIGURE 16A

9961	AAAAAAAAA AAAAA
1950	TITGITITGI TATITGCAGI ITACAAATAT AGTATTATIC TCTAAAAAA
0061	ATACTTCATG TGTAATTATA GCTTAGACTT TAGCCTTCTT GGACTTCTGT
1850	ACATCTGGAT ATCTTGTCAC ATTTTGTAC ATTGTGACTG CTTTCAACAT
0081	ACCAGATGAA ACTGGATATA ATTTGAGACA AACAGGATGT GTTTTTTTAA
1750	TTACCAGCAG GAGGGAAACA CACTTCACAC AACAGGCTTA TCAGAAACCT
0001	CCAAGTAGAC TCACTCCTTA AAAAATTTAT GGATATCTAC AAGCTGCTTA
16 50	CAGTITICAGA CGAATGAATG TAATAGGAAA CTATGGAGTT ACCAATATTG
1600	TGAAAAACTG GCAAAGGAAT GGAAGAATAT TTTAGTCATG AAGTTGTTTT
1550	TGTTGATGAC AAGAAGCAAT ACATTCCAGC TTTTCCTTTG ATTTTATACT
1500	CTTGACCAAA GAACACTTGA TTAGGTTAAT TACCCTCTTT TTGAAATGCC
1450	AAAAAAGGA TTATGAAAGG ATTTAGAACA AAGTGGAAGC CTGTTACTAA
1400	GCCAGAAAAG AGGCTGTTTG TAAGATTCAG GCCATACTGG AAAAATTAGA
1350	TITTGGAACT GGATTCAGTT GAAACTGGGG GCCAGGACTC TGTACGGCAG

FIGURE 16B

DAWASPGAYGMGGRYPWPSSAPSAPPGNLYMTESTSPWPSSGSPQSPPSPPVQQPKDSSYPYSQSDQSMNRHNFPCSVHQ EPGRAGGSHQEQPPYPSYNSNYWNSTARSRAPYPSTYPVRPELQGQSLNSYTNGAYGPTYPPGPGANTASYSGAYYAPGY TQTSYSTEVPSTYRSSGNSPTPVSRWIYPQQDCQTEAPPLRGQVPGYPPSQNPGMTLPHYPYGDGNRSVPQSGPTVRPQE MSALARSGYGPSDGPSYGRYYGPGGGDVPVHPPPPLYPLAPEPPQPPISWRVRGGGPAETTWLGEGGGGDGYYPSGGAWP YESSGTVINEDSDLLDSQVQYSAEPQLYGNATSDHPNNQDQSSSLPEECVPSDESTPPSIKKIIHVLEKVQYLEQEVEEF VGKKTDKAYWLLEEMLTKELLELDSVETGGQDSVRQARKEAVCKIQAILEKLEKKGL

FIGURE 16C

0	CIC	XXC.	LOC (x	icox	CA.A	ca	CT	TCA	∞ c	CM	KCC.	-X TC																r s	cc y
																													,crocc	<u>c1</u> 79
•	G	ĸ	Y	¥	C	P	G	Ç	C	Ð	٧	₽	V	H	P	P	₽	P	L	Y	*	۲	Ŗ	₽	Ε	P	P	O	P	
~	~ . 1		٠,	~~	700	rcc	:ca	200	cca	<u></u>	000		دعد	ىص	œ	TOCK	76	227	C	220	_	æ	ιœ	CAI	900		TA I	ccc	1000	GA 27
- P	ĭ	5	H	×	٧	R	G	G	G	P	A	£	T	T	w	L.	G	Z	G	G	c	c	D	0	Y	Y	P	s	G	
																													TCTA	CI 36
3	λ	H	r	E	P	a	R	λ	G	a	3	H	Q	£	Q	P	P	Y	P	5	Y	N	8	N	Y	W	H	2	Ŧ	
~	^-		~~.	~~~		~~~	~~	~	C= 1	~ 1	. TY	·	TA B	GAC	~L	38 R 3	~~ C2	~ .	~~	-10	A COT		- 2 2 7	~~	~~ A T	'A/"B		50	CCT.	AT 45
						¥																								
•	•	-	•	••	•	-	•	_	-								_		_	-	_	•	_	-	-					
																													TACT	CC 54
3	P	T	Y	r	r	C	P	G	٨	24	Ŧ	λ	5	Y	5	a	¥	Y	Y	λ	P	G	¥	T	Q	T	S	¥	6	
						~~.			~.~	~	~~		cor	~	<u>_</u>	~10		~~~	~	~~	. ~~		~~	~**	~	~10	****	~2	acts.	n a
						Y																								
•	_	Ť	•	-	•	-	•	-	_			_					-	•	••	_	-	•	_	-	_	_	-	•		
																													CCIA	AT 72
N.	P	P	L	ĸ	C	Q	v	₽	G	Y	P	P	3	Q	×	•	G	¥	T	I.	P	Œ	¥	P	Y	G	Ð	C	¥	
_				.		~	~	~~				~~	- 2 2 6			~ ~~	~~	~	~~~	~~	~~~	~~	 - • •			~~	~~~	~~	TATO	CC EI
						~																								-
	_			_								_									-	-								
																													CAGT	CA 90
K	P	S	s	Х	P	z	٨	P	r	٥	24	Ľ	¥	M	T	Z.	s	T	2	P	W	r	s	s	G	S	₽	Q	2	
~	~		~•~	~~		~~ ·	. ~~	# CC	~~		2 Total		TAT	· » ~	٠,	Ta	c	- A A -	TC & C	: እ. ተ/	-11	.	· 1 TY		***	-	***		CCIN	GC 99
						0																								
			-			_	•													_					•					
																													ccic	AG 108
2	*	R	Q	Y	Ε	2	Ξ	C	T	V	N	N	Z	Þ	S	D	L	L	D	5	Q	V	Q	r	\$	X	E	P	Q	
~	Y-T	. ~~	CT1	N TYC	~~	~~		E (7)	E 777	~~ 1	1 C1	.a TY		A TY	`A A I	CT.	GC.		~	~~~		~		Y-74	~	-TC-3		~~	AGTA	CT 117
						s																								_
																													ACAG	AC 126
₽	r	3	1	K	K	I	I	ĸ	v	L	E	K	V	0	Y	L	E	Q	ĸ	V	E	E	2	v	G	K	K	T	ь	
A J	LAG	CAT	A C-1	<u> </u>	111	TOC.	. LC		TOC	TAA	בכא	J.GC	i	TTI	TC	3 3.20	TO	ea Tr	r ac	- - -		ac-t	~~	oα	CAG	car.c	701	GD.	vecce	AG 135
						E																								
																						TC.	w	GAT	TTA	OV.	(C)	NG.	CCAN	GC 144
×	R	K	E	^	V	C	K	I	Q	λ	I	L	E	K	L	Ε	K	K	C	L	•									
<u>-</u> 2	G.	D.C	TAR	CIN	20	200	ىت	C -4	cric	AT	cua		ART	TAC	œ	CIT		3 11	ATG	œ		CA:	CA			œ.	ATAC	AT	rccus	c 15:
																													LTGLA	
																													CACC	a 17
																													1111	
																				LTO	tot	w	TA1	144	111	av		NC(CLIC	7 189 19
•	~=/		CTO	777		TTT	777	17	TOO	VGT.	TA	TA.A	ATX	TAC	TAT	TAT	CT	~												

CCCCCCCCC CCCCCCCC CCNGAAGACG CCCGGAGCGG CTGCTGCAGC 50	20
CAGTAGCGGC CCCTTCACCG GCTGCCCCGC TCAGACCTAG TCGGGAGGGG 10	8
TGCGAGGCAT GCAGCTGGGG GCCCAGCTCC GGTGCCGCAC CCCGTAAAGG 16	150
	200
	250
ATATGGGAAA CCAACATCCT TCTATTAGTA GGCTTCAGGA AATCCAAAAG	300
GAAGTAAAAA GTGTAGAACA GCAAGTTATC GGCTTCAGTG GTCTGTCAGA	350
TGACAAGAAT TACAAGAAAC TGGAGAGGAT TCTAACAAAA CAGCTTTTTG	400
AAATAGACTC TGTAGATACT GAAGGAAAAG GAGATATTCA GCAAGCTAGG	450
_	200
	550
	009
	650
	700
4	750
ထ	800
	850
	006
	950
	1000
	1050
	1100
.	1150
•	1200
	1250
AGTGATCGAG GTGCAAACTC TGATCACATA TATTGACTTG AAGGAGGCCC	1300

α	1350
	1400
	1450
	1500
L	1550
ග	1600
ATCTCACTTT TGATACTGTT TTGCACTTCA TATGTGCTTC TATGTATAGA 1650	20
	00
TGATTGAAGC AAATTCTATT CAGTATCTGC TGCTTTTGAT GTTGCAAGAC 1750	,50
ပ္	1800
ATGTGGTGTG GTTTGTTTGG TTTGTCCTTT TTTTTGCGTT TTTAATCAGA 1850	0
-	1900
	1950
 - -	2000
	2050
TITITICACT TATAAGGCCT GATTGGTCCT ACCCAGCTTA ACGGGGTGGG 2100	00
	0
	003
	2250
 	2300
TGTAAGTTGC TTTTGTTTGT AAAATCTGAG CTTATAGTTT TCCTTAGTGA 2350	20
	2400
	2450
(D	2500
TAGCACATAA TGGTTCTCTT CTGTTGTCCA AGGCTGTAAA ATGGACAGCC 2550	50
TTGTCACACC TCCCCGGTGC TGTTTTACAA CGTGAGGGTA GACGCTGTCA 2600	90

GTAACCCAGA GGGACCAGGC CTTCCTAGGT TTTCTAGGCA GTCAGCTGTT	2650
_	2700
GTGAAACCTG CTCGGAATTA AAGGCTTCCT CTGGGTGCCT GCTGAACAAC	2750
\circ	2800
ATTAAAGTCA GTCGTGCGTG AAGCATCTCT CTTCTAAAGG ATGTGTATTT	2850
CCATAAATGC TTTCTGAGGA TCCGGTACAA AATGATTTCC CAAAGTTCTG	2900
AAGTGCCTTG AGAACATGTG GGTCCGAGTG TTATAACAGA CTCCTCCCCC	2950
GGGTCACCTT TTGCCTGGTC ATCCTGTTAG AGTACATCTT TGGAAATCCA	3000
GGGTAATATT CTCTTTCAGA GATGCTCATT GTGTAACTCT GTGTAGGGAG	3050
ATAGTCACTT TAAACAGCTC AAAGTAGCTA GCTAAAGGAG TAGCCTTAAA	3100
TACCTAAAAG ATGACAGAAG CATAGCCCTT AACAAATCTT CAGCTTGTCT	3150
	3200
	3250
	3300
TCAAGGCAGG TCATTGGAAT CCACGTTTTG GCCACAGTAG TTGTAGGATT	3350
GCTTTTCTGT ATCATAATTT TAGAATGCTC TTAAAATCTT GAGGAAGAGT	3400
	3450
	3500
۲ŋ	3550
	3600
	3650
GOCCCCCAAA GTGCTGGGAT TAACGGGTGT GAGCCACGGC GCCCAGCCCA	3700
GGAAGAGTTT TTAAATTAGA GCTCTGTTTA ATTATACCAC TGGGAAATCA	3750
	3800
TTTCATTTTG TAAAGTTAAA TGTCAGCATT CCCTTTAAAA GTGTCCATTG	3850
TTCTTTGAAA GTAGACGTTT CAGTCATTCT TTTCAAACAA GTGTTTGTGT	3900

3950 4000 4050 4100 4200 4250 4300 GTATTTTGT GATCTGTAAT GAAAAGAATC TGTACTGCAA GTAAAACCTA CTCCCCAAAA ATGTGTGGCT TTGGGTCTGC ATTAAACGCT GTAGTCCATG TCAGATTGAC CTTGATTGAC TGTCAGGCAT GGCTTTGTTT CTAGTTTCAA AATATGAGCT ACTGCATGTA ATTCTTAAAC TGGGCTTGTC ACATTGTATT ACCTITIGCC AAGCIGIGG CATCGIGIT GAGTACAGGG IGCICAGCIC TTCCACCGTC ATTITGAATT GTTCACATGG GTAATTGGTC ATGGAAATGA TCTGTTCTCG TTCCTTGTAC CGGATTATTC TACTCCTGCA ATGAACCCTG TTGACACCGG ATTTAGCTCT TGTCGGCCTT CGTGGGGAGC TGTTTGTGT **ITCATGCC**

FIGURE 17B

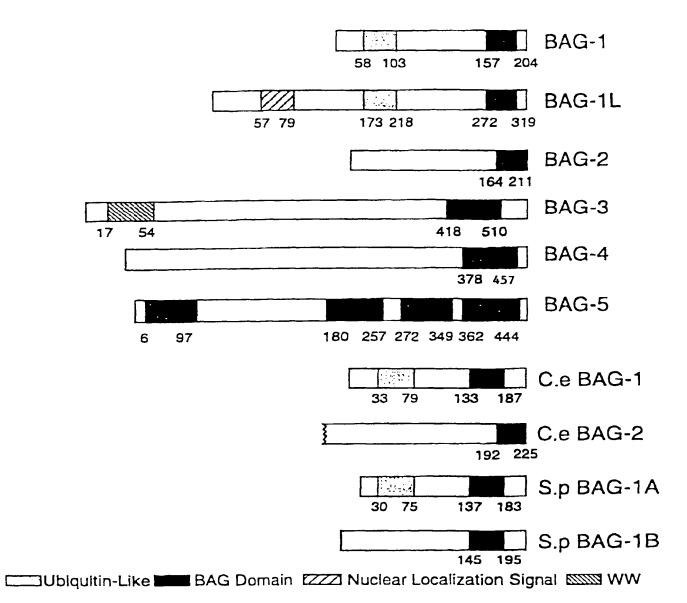
400 447 200 250 300 100 ALEKRKLFAC EEHPSHKAVW NVLGNLSEIQ GEVLSFDGNR TDKNYIRLEE IALLMGVNNN ETCRHLSCVL SGLIADLDAL DVCGRTEIRN YRREVVEDIN KLLKYLDLEE EADTTKAFDL RQNHSILKIE KVLKRMREIK NELLQAQNPS HTLTKICAVQ EIIEDCMKKQ PSLPLSEDAH PSVAKINFVM CEVNKARGVL MDMGNQHPSI SRLQEIQKEV KSVEQQVIGF SGLSDDKNYK KLERILTKQL LLTKQLLALD AVDPQGEEKC KAARKQAVRL AQNILSYLDL KSDEWEY FEIDSVDTEG KGDIQQARKR AAQETERLLK ELEQNANHPH RIEIQNIFEE AQSLVREKIV PFYNGGNCVT DEFEEGIQDI ILRLTHVKTG GKISLRKARY ELYLSSKTEL QGLIGQLDEV SLEKNPCIRE ARRRAVIEVQ TLITYIDLKE

34 / 35

FIGURE 17C

coccoccc	∞	CONCARCACO	COOLGACCEC	CTCCTCCACC	CHCTHCCCCC	COCCTTONCOC	ectrococc	TOAGROCTING	90
TOCCCACCCC	TGCGAGGCAT	CONCETCOCC	SCOCASCTCC	CCTOCCCCCAC	CCCCTHAAGC	OCTORICTIC	CHOCTOSCOA	OCTORCOORC	140
CCCACGCCAA	CACCECATCC	ANTTCREACT	*CTTTTCTC	CTTCTCAAAC	TGRACACRAC	AAAACTRTCC	NTRTGGGAAA	CCAACATCCT	270
TCTRTTRCTR	ESCTTCASSA L Q I	I Q K	CAACTRAAAA I V K J			CCTTCACTC	ETCTCTCACA	TCACAGGAAT B K N	340
THOMCANAC T K K L	TOGACAGGAT I R I	TCTRACAAAA L T K	d r i i			ONACCAMANC E & & &	CACATRITICA P 1 C	COAACCERCC 4 A L	450
ANOCCCCCAC K R A A	CHCHCGAGAC	ACAMOSTOTT E R L	CTOMMENCT L K E L			CONCACCOCA 2 E I I	TTORRATROR I I Q	H I I	£40
ERGGRAGECC I I A Q	ACTOCCTOCT S L V	ERCACACAAA 1. I K	ATTGTGCCAT			CDIACTORTS	ACTTICAACA I I I	e I q	430
CATRICATIC	TGAGGCTGAC R L T	ACATOTINAA K V K	ACTGGAGGAA T G G K			ACCTRICACA R Y E T	CTTTMACCAA L T K	AATCTETECE 1 C A	720
CTGCAAGAGA	TRATCGAAGA	CTCCATGAAA	AACCACCCTT K Q 2 3	L I L		SCACATCCTT A K I I	CCCTTGCCAA	ARTCHACTIC	410
T H C I	ACCTGAACAA V H K		A T I W			RACRATGAGA N N I T	C R E	CTRICCICI L J C	900
A T 3 e	L I A		ecicarcate			AGAAATURTC R K T R	CCACCCACCT R I V	ACTRICARGET V I 3	990
ATCHACHART I N K L	ERTICAMETR L K I	TCTCCATTIC	CAACACCAAC I I I A			SACCTGAGAC B L 1 Q	TITACTARGA L B H	CRTITINAAA 1 L K	1000
ATRGAAAAGG I I K 9	TOUTOMAGNG		RTRAGGATT I K H I			CCTICTCAAT 7 S I L	TUTHOCTORC Y L S	CTCCRAAACR J K T	1170
CAATTCCACC	L I 6	ACACTICGAT		TICAAAAAA E K H		CCCCAACCCA R I A R		ACTGATOGAG V I I	1240
T Q T L	TGATCACATA I T I	TATTCACTTC I B L				A C I I	MCCACCCATC EL 2 3	E K K	1350
A A N A	TCCTTCGAAA	CTTCTCTCAG L J I	ATCCAGGGAG I Q G I			AATCGAACCG H X T 3	ATRAGAACTR K H Y	I R L	1440
EARGACCTCC I I L L	TCACCAAGCA T K Q	E L A	CTGGATGCTG L 3 A V			MACTOTRACE K C K R	A R K	RCAAGCTGTG Q R V	1230
ACCCTTGCGG	AGAATRITUT H I L		SACCTGARAT B L K S	CYCATGAATG	CCACTACTGA I Y .	MATRICONGRG	ATCTCACTTT	TOATHCTGTT	1620
TTGCACTTCA	BACTCCTTC	TRICIATACA	CACCTITOAC	TTCATTCATT	TRITROCTCON	DITTTCACTO	TORGERITIES	TCATTCAACC	1710
**************************************	CACTRICICC	TOCTTTICAT	CTTOCAAGAC	MANDRICATT	ACAGCACCTT	MACTITICOA	TTOCCATORY	MICICINIC	1600
#1c1cc1c1c	CTTTCTTTCC	TITGECTAT	TITTICCOTT	TIDANICACA	MACAAARTR	CACCCACCTT	TICTACRITI	TMATCCCTT	1690
CTCCRACCAT	TRANSFERE	STCTTTOACA	ATCTRGAACT	AGOCATRACC	TIRCATRATA	CENGGAAAAT	TRTCACAAAC	CCCAMATITI	1440
	ACACTRACCT								2070
CATTOCTOCT	ACCCAGCTER	ACCCCCTCCC		TTETTCACAC	ACTOTOTICT	TTTTTTTAACA	11111HC11C	CARARACAGE	2140
ATCTGCATTT	TOOCCHTOCT	CDUCCILIAN	CACACCAATC	*********	CTCCMACATR	RCARARTTRE	CANACTIVIT	ACCOPARAMA	2,250
	TOCHTHAG	-							2340
	CHANTICHT								2430
	REDICTORNOR								2520
	ACCUTURAN								2410
	CLICCINCEL								2790
	CTOCCAATTA								2460
	CHARGETON								2970
	ACTACATCTI								3060
	MACTROCTA								3160
CTONGERTY	COOMITORIS	*******************************	SCHOOL STATE	TOUTHOR	MICTICIAL	MATCOCTOOR	COCTOCCCTC	ROACCCCACT	3240
COCINITION	CODETCACCE	TOCONTRCER	CTOOMSCTCC	MCTCCTCCC	*************	TOMOCOCHOS	TORTTOGART	CONCETTIC	3330
	TICTROCATE								34 2 0
	rereserve:								36 10
TOCCTOROCO	ACCTGACTAG	CTGGGACTRC	ACCOUNTCICS	CHCCHTCCCT	COCHMITTI	TURNITION	ATHEACTTON	SATTTORCOA	3400
	. ectectere								34+0
COCACCCO	N DCAAGACTTT	TEMATTRICA	OCTOTOTAL	RTTHTRCCAC	TCCCAAATCA	TOGTTMOSCT	TORCCONTRI	ECTTOCOCAG	2760
								CERCACCITY	3670
CACTCATTC	I TITCAAACM	etettietet	**CCTTTICCC	ANCETETECE	CATCETETET	CACTROACGE	TOCTOMOCYC	TTOCACCETC	3140
								CENCTITOM	
								CCTOCOCACC	4140
								EARAGERETC	Q 30
TUDACTOON	N CENAMACCIA	CTCCCC	ATCTCTCCCT	**************************************	RTTNAACGCT	CTMCTOCATG	TTOATGOC		4304

FIGURE 18



SEQUENCE LISTING

<110>	Reed, C	John	С.												
	Takayan	nā, S	Shin.	ichi											
	The Bur	rnhai	m In.	stit	ute										
<120>	Novel E	BAG :	Prot	eins	and	Nuc	leic	Aci	d Mo	lecu	les	Enco	ding		
<130>	FP-LJ	3646													
<140>															
<141>															
<150>	09/150	, 489													
<151>	1998-0	9-09													
<160>	24														
<170>	Patent	In V	er.	2.0											
<210>	1														
<211>	1291														
<212>	DNA														
<213>	Homo s	apie	ns												
		•													
<220>															
<221>	CDS														
	(46)	(108	0)												
	, ,	,	•												
<400>	1														
	geget e	aget	tcca	it co	ictac	acad	ı tca	acaa	agtg	caac	ac ct	a qo	et da	ag cgc	57
	J - J	3		-		, , , ,			5 5	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,				ln Arg	
												1		,	
aaa a	gg gcg	caa	aga	ccq	cqa	ggc	gac	cgg	gag	cgq	ctg	ggt	tcc	cgg	105
	ly Ala														
5		5	5	10		_	•	_	15			-		20	
_															
ctac	gc gcc	ctt	caa	cca	aac	caa	gag	cca	cac	cag	tca	gag	ccc	cca	153
	rg Ala														
1 100 1	119 1114	пса	25	110	O ± y	9	010	30	•••	011.			35		
			23					0					0.0		
000	ag cgt	aat	cca	cct	CCC	t C t	רממ	cat	CC3	cct	acc	caa	ant	act	201
	ag egt Sln Arg														201
MIG G	in Arg	40	LIO	FIO	-10	Set	45	nt à	110	FIO	LTQ	50	261	1111	
		4 ()					4.7					50			
		A = E	~	0.5	0.00	200	200	ac.c	a cc	acc	900	999	act	606	249
gcc a	igc ggg	cat	gac	cga	CCC	acc	agg	ggc	gee	gcc	gee	gge	get	age	249

Ala	Ser	Gly 55	His	Asp	Arg	Pro	Thr 60	Arg	Gly	Ala	Ala	Ala 65	Gly	Ala	Arg	
	_		_	-	-			cgg Arg	-	-	-			-		297
	_			_			-	acc Thr	_	_		_				345
=	-	- •			_	_		gag Glu			_		_		-	393
		_	-					gac Asp 125			_			-		441
					-	-		gca Ala	-				-		_	4 89
		-			-		-	ctt Leu		-			_	_		537
								ctg Leu								585
		-		_			_	aaa Lys				_				633
_	_	-	-	•		-	-	tca Ser 205	-					-		681
_		_	_					aag Lys		_		_	-		-	729
				_				gag Glu			_	-				7 77
gac	cag	ctg	gaa	gag	ttg	aat	aaa	gag	ctt	act	gga	atc	cag	cag	ggt	825

Asp 245	Gln	Leu	Glu	Glu	Leu 250	Asn	Lys	Glu	Leu	Thr 255	Gly	Ile	Gln	Gln	Gly 260	
	ctg Leu		_	_			•	•	•		•			-		873
_	gta Val		_			-										921
	aca Thr			_		_				-	-	-	_			969
	ggc Gly 310	_	-		_	_	_	_			_	_ ,	-	_		1017
	gag Glu	_			-	_			-	_						1065
	gcc Ala				tga	ggtg [,]	tag	caga	aaaa	gg c	tgtg	ctgc	c ct	gaagi	aatg	1120
gcg	ccac	cag	ctct	gccg	tc t	ctgg	atcg	g aa	ttta	cctg	att:	tott	cag	ggct	gctggg	1180
ggc	aact	ggc	catt	tgçc.	aa t	tttc	ctac	t ct	caca	ctgg	ttc	tcaa	tga	aaaa	tagtgt	1240
ctt	tgtg	att	tgag	taaa	gc t	ccta	ttct	g tt	tttc	acaa	aaa	aaaa	aaa	a		1291

<210> 2

<211> 345

<212> PRT

<213> Homo sapiens

<400> 2

Leu Ala Gln Arg Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg 1 5 10 15

Leu Gly Ser Arg Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln
20 25 30

Ser Glu Pro Pro Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro 35 40 45

Ala Arg Ser Thr Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala Ala Gly Ala Arg Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser Thr Arg Ser Glu Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu Glu Ala Thr Trp Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln Gly Glu Glu Met Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser Thr Arg Ser Glu Glu Val Thr Arg Glu Glu Met Ala Ala Gly Leu Thr Val Thr Val Thr His Ser Asn Glu Lys His Asp Leu His Val Thr Ser Gln Gln Gly Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val Val Glu Glu Val Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe Lys Gly Lys Ser Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly Ile Gln Asp Gly Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro Gln Glu Glu Val Glu Leu Lys Lys Leu Lys His Leu Glu Lys Ser Val Glu Lys Ile Ala Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly Ile Gln Gln Gly Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys Lys Leu Asp Arg Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile Leu Glu Glu Ile Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser

Arg Leu Lys Arg Lys Gly Leu Val Lys Lys Val Gln Ala Phe Leu Ala 315 365 310 Glu Cys Asp Thr Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu 325 330 Gln Ser Thr Asn Phe Ala Leu Ala Glu 345 340 <210> 3 <211> 1179 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (160)..(792) <400> 3 geageegegg tgtegegaag teeteeeggg ttgeeeeege ggegteagag ggagggeggg 60 egeogegttg gtgaeggega ceetgeagee caaggagege teeacteget geogeeggag 120 ggccggtgac ctcttggcta ccccgcgtcg gaggcttag atg gct cag gcg aag 174 Met Ala Gln Ala Lys 1 ate aac get aaa gee aac gag ggg ege tte tge ege tee tee tee atg Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys Arg Ser Ser Met 10 15 20 270 get gac ege tee age ege etg etg gag age etg gae eag etg gag ete Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu Asp Gln Leu Glu Leu 30 35 25 agg gtt gaa gct ttg aga gaa gca gca act gct gtt gag caa gag aaa 318 Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala Val Glu Gln Glu Lys 45 40 gaa atc ctt ctg gaa atg atc cac agt atc caa aat agc cag gac atg 366 Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln Asn Ser Gln Asp Met 55 60 agg cag atc agt gac gga gaa aga gaa tta aat ctg act gca aac 414 Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu Asn Leu Thr Ala Asn 80 85 70 75

														aca Thr 100		462
-														att Ile		510
-			_											aag Lys		558
		_												cat		60h
	-	_												ctt Leu		654
-	_													aat Asn 180		700
_														gga Gly		750
					caa Gln											793
tagi	tctt	caa .	acct	aaga	gc a	ttta	caca	a ta	caca	aggt	gta	aaaa	tga	taaa	atacta	852
ttt	taat	tga	taac	tagt	tc t	ttgt	tagg	t at	aacc	actt	agt	tgac	act	gata	gttgtt	912
tca	gatg	agg	aaaa	tatt	cc a	tcaa	gtat	c tt	cagt	tttg	tga	ataa	caa	aact	agcaat	972
att:	ttaa	tta	tcta	tcta	ga g	attt	ttta	g at	tgaa	ttct	tgt	cttg	tac	tagg	atctag	1032
cat	attt	cac	tatt	ctgt	gg a	tgaa	taca	t ag	tttg	tggg	gaa	aaca	aac	gttc	agctag	1092
9 99	caaa	aag	catg	actg	ct t	tttc	ctgt	c tg	gcat	ggaa	tca	cgca	gtc	acct	tgggca	1152
ttt	agtt	tac	taga	aatt	ct t	tact	gg									1179

<210> 4

<211> 211

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Gln Ala Lys Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys
1 5 10 15

Arg Ser Ser Met Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu 20 25 30

Asp Gln Leu Glu Leu Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala 35 40 45

Val Glu Gln Glu Lys Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln 50 55 60

Asn Ser Gln Asp Met Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu 65 70 75 80

Asn Leu Thr Ala Asn Arg Leu Met Gly Arg Thr Leu Thr Val Glu Val
85 90 95

Ser Val Glu Thr Ile Arg Asn Pro Gln Gln Gln Glu Ser Leu Lys His 100 105 110

Ala Thr Arg Ile Ile Asp Glu Val Val Asn Lys Phe Leu Asp Asp Leu 115 120 125

Gly Asn Ala Lys Ser His Leu Met Ser Leu Tyr Ser Ala Cys Ser Ser 130 135 140

Glu Val Pro His Gly Pro Val Asp Gln Lys Phe Gln Ser Ile Val Ile 145 150 155 160

Gly Cys Ala Leu Glu Asp Gln Lys Lys Ile Lys Arg Arg Leu Glu Thr 165 170 175

Leu Leu Arg Asn Ile Glu Asn Ser Asp Lys Ala Ile Lys Leu Leu Glu 180 185 190

His Ser Lys Gly Ala Gly Ser Lys Thr Leu Gln Gln Asn Ala Glu Ser 195 200 205

Arg Phe Asn 210

<210> 5 <211> 2528 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (1)..(2031) <400> 5 geg gag etc ege ate caa eec egg gee geg gee aac tte tet gga etg 48 Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Asn Phe Ser Gly Leu 1 5 10 gac cag aag ttt cta gcc ggc cag ttq cta cct ccc ttt atc tcc tcc 96 Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser 20 25 ttc ccc tct ggc agc gag gct att tcc aga cac ttc cac ccc tct Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser 35 40 ctg gcc acg tca ccc ccg cct tta att cat aaa ggt gcc cqq cqc cqq 192 Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg 55 ctt ccc gga cac gtc ggc ggc gga gag ggg ccc acg gcg gcc cgg 240 Leu Pro Gly His Val Gly Gly Gly Gly Pro Thr Ala Ala Arq 65 70 75 cca gag act egg ege eeg gag eea geg eee ege ace ege gee eea geg 288 Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala 85 90 ggc aga ecc caa ecc age atg age gec gec acc cae teg ecc atg atg 336 Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met 100 105 cag gtg gcg tcc ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg 384 Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp 115 120 125 gag atc aag atc gac ccg cag acc ggc tgg ccc ttc ttc gtg gac cac 432 Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His 135 140 aac age ege ace act acg tgg aac gae eeg ege gtg eee tet gag gge 480

Asn 145	Ser	Arg	Thr	Thr	Thr 150	Trp	Asn	Asp	Pro	Arg 155	Val	Pro	Ser	Glu	Gly 160	
				cca Pro 165												528
	_	-		gct Ala		-				- •			_		-	5 76
				ccc Pro												624
				ttc Phe		-			_			_	_	_		672
_		_		gca Ala												720
				gaa Glu 245												768
		-		gca Ala												816
	_			gct Ala	_		-	_							-	864
	_			tcc Ser												912
	, , ,			tcc Ser		-							-			960
	_	-	_	ccc Pro 325												1008
gcg	cag	agg	ggt	gag	tac	cag	acc	cac	cag	cct	gtg	tac	cac	aag	atc	1056

RNSDOCIDE ZWO D01410641 I S

Ala	Gln	Arg	Gly 340	Glu	Tyr	Gln	Thr	His 345	Gln	Pro	Val	Tyr	His 350	Lys	Ile	
		,	-	,,					-			-	tcc Ser	_		1104
													cca Pro			1152
_	-	=						_			-		cac His			1200
										_	_		gca Ala		_	1248
											-		gtt Val 430			1296
										-		-	aaa Lys		-	1344
_				-		_	-						gag Glu	_	_	1392
						-		_		-			ccc Pro	_		1440
											-	-	aca Thr	-		1488
				_				-	-				cca Pro 510			1536
	-	-		-								-	aaa Lys		-	1584
gcc	atc	ctg	gag	aag	gtg	cag	ggg	ctg	gag	cag	gct	gta	gac	aac	ttt	1632

Ala	Ile 530	Leu	Glu	Lys	Val	Gln 535	Gly	Leu	Glu	Gln	Ala 540	Val	Asp	Asn	Phe	
-		-	-		-		-		-	-		_	, ,	tat Tyr		1680
														cga Arg 575	-	1728
													_	acc Thr		1776
										-				gtc Val	_	1824
-		-		-		_			_	-	_	_		ctg Leu	-	1872
-		_		-		-		-	-	_	-		_	aaa Lys		1920
														gaa Glu 655		1968
	_	_						_	-	_		-		cct Pro	ggt Gly	2016
	cca Pro				tag	cctc	tgc (cctg	taaa	ag to	cagad	ctcg	g aa	ccgat	tgtg	2071
tgc	ttta	ggg (attt:	tagt	tg c	atgc	attt	c aga	agac	ttta	ggt	cagt	tgg '	tttt	gattag	2131
ctg	cttg	gta	tgca	gtac	tt g	ggtg	aggca	a aa	cact	ataa	agg	gcta	aaa (ggga	aaatga	2191
tgc	tttt	ctt	caat	attc	tt a	ctct	tgta	c aa	ttaa	ngaa	gtt	gctt	gtt	gttt	gagaag	2251
ttt	aacc	ccg	ttgc	ttgt	tc t	gcag	ccct	g tc	nact	tggg	cac	cccc	acc .	acct	gttagc	2311
tgt	ggtt	gtg (cact	gtct	tt t	gtag	ctct	g ga	ctgg	aggg	gtad	gatg	3 99	agtc	aattac	2371

ccatcacata aatatgaaac atttatcaga aatgttgcca ttttaatgag atgatttct 2431 tcatctcata attaaaatac ctgactttag agagagtaaa atgtgccagg agccatagga 2491

atatotgtat gttggatgao tttaatgota catttth

2528

<210> 6

<211> 677

<212> PRT

<213> Homo sapiens

<400> 6

Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Ala Asn Phe Ser Gly Leu
1 5 10 15

Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser 20 25 30

Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser 35 40 45

Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg 50 55 60

Leu Pro Gly His Val Gly Gly Glu Gly Pro Thr Ala Ala Arg
65 70 75 80

Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala 85 90 95

Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met 100 105 110

Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp 115 120 125

Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His 130 135 140

Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly 145 150 155 160

Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser 165 170 175

Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg

185 190 180 Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg 205 195 200 Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe 220 210 215 Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu 235 230 225 Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val 250 255 245 Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg 265 270 260

Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ser Ala 275 280 285

Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro 290 295 300

Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg 305 310 315 320

Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro 325 330 335

Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile 340 345 350

Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe 355 360 365

Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg 370 375 380

Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val 385 390 395 400

Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val 405 410 415

Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro 420 425 430

Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val

Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly

Asn Pro Ala Ala Pro

<211> 1010 <212 DNA <213> Homo sapiens <220> <221 > CDS <2223 (323)..(1009) <400> 7 acquitateet qtaaqaccaa qaattgcaaq gecaqaqttt gaattettat acaaatggag 60 egtatggtee aacatacece ecaggeeetg gggeaaatae tgeeteatae teaggggett 120 attargeace tggttatact cagaccagtt actccacaga agttccaagt acttaccgtt 180 catciggeaa cageecaact ceagtetete gitiggateta teeccageag gaetgicaag 240 actgaagcac cocctettaa ggggcaggtt ccaggatate egeetteaca gaaccetgga 300 atgaeectige decattated to atg gag atg gta atc gta gtg toe cac aat Met Glu Met Val Ile Val Val Phe His Asn cae qge eqa etq tae qae cae aaq aaa qat geg tqq get tet eet ggt His Gly Arg Leu Tyr Asp His Lys Lys Asp Ala Trp Ala Ser Pro Gly 15 20 25 get tat gga atg ggt gge egt tat eee tgg eet tea tea geg eee tea 448 Ala Tyr Gly Met Gly Gly Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser 30 35 gea eea eee gge aat ete tae atg act gaa agt act tea eea tgg eet 496 Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro 45 50 age agt gge tot occ dag toa occ oct toa occ doa gto dag dag occ Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val Gln Pro 60 65 70 aag gat tot toa tao coo tat ago caa toa gat caa ago atg aac egg 592 Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg 80 75 85 cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tcg ggg aca gtg His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser Gly Thr Val

100

105

95

<210: 7

	gat Asp								688
	cag Gln 125								736
-	agt Ser								784
	ccg Pro	-							832
	gaa Glu								880
_	tgg Trp								928
	gtt Val 205								976
	gtt Val					a			1010

<210> 8

<211> 229

<212> PRT

<213> Homo sapiens

<400> 8

BNSD00ID <WC __0014106A1 . >

Met Glu Met Val Ile Val Val Phe His Asn His Gly Arg Leu Tyr Asp 1 5 10 15

His Lys Lys Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly
20 25 30

Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu 35 40 45

Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln
50 55 60

Ser Pro Pro Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro 65 70 75 80

Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser 85 90 95

Val His Gln Tyr Glu Ser Ser Gly Thr Val Asn Asn Asp Asp Ser Asp 100 105 110

Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly
115 120 125

Asn Ala Thr Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Leu 130 135 140

Lys Ile Ile His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val 165 170 175

Glu Glu Phe Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu 180 185 190

Glu Met Leu Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly
195 200 205

Gly Gln Asp Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile 210 215 220

Gln Ala Ile Leu Glu 225

<210> 9

<211> 689

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (3)..(482)

<220>

<221> unsure

WO 00/14100		101,00,0,210
<222> (105) <223> any amino acid		
• •	aa ctt ctc caa gca caa aac cct lu Leu Leu Gln Ala Gln Asn Pro 5 10	
3 3	aca gaa ttg cag ggt tta att gg Thr Glu Leu Gln Gly Leu Ile Gl 25	3 3 3
	aaa aac ccc tgc atc cgg gaa gc Lys Asn Pro Cys Ile Arg Glu Al 40	
	caa act ctg atc aca tat att ga Gln Thr Leu Ile Thr Tyr Ile As 55 6	, , , ,
	aag ctg ttt gct tgt gag gag ca Lys Leu Phe Ala Cys Glu Glu Hi 70 75	
	gtc ctt gga aac ttg tct gag at Val Leu Gly Asn Leu Ser Glu Il 85 90	
	gga aat cga acc gat aag aac ta Gly Asn Arg Thr Asp Lys Asn Ty 105	
	aag cag ctg cta gcc ctg gat gc Lys Gln Leu Leu Ala Leu Asp Al 120	
	tgt aag get gee agg aaa eaa ge Cys Lys Ala Ala Arg Lys Gln Al 135 14	a Val Arg Leu
	agc tat ctc gac ctg aaa tct ga Ser Tyr Leu Asp Leu Lys Ser As 150 155	

18

tac tgaaatacca gagatctcac ttttgatact gttttgcact tcatatgtgc

532

Tyr 160

ttotatgtat agagagettt eagtteattg atttataegt geatatttea gteteagtat 592
ttatgattga ageaaattet atteagtate tgetgetttt gatgttgeaa gaeaaatate 652
attacageae gttaaetttt eeatteggat eaaaaaa 689

<210> 10

<211> 160

<212> PRT

<213> Homo sapiens

<400> 10

Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr 1 5 10 15

Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu
20 25 30

Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Ala 35 40 45

Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala 50 55 60

Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys 65 70 75 80

Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val 85 90 95

Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu 100 105 110

Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln 115 120 125

Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala 130 135 140

<210> 11

<211> 246

<212> DNA

<213> Caenorhabditis elegans

<400> 11
atgtetttee geetettegt tgaaatattt cactttettt teeagetttt teeceatete 60
gaeetgettt ggtttttega gaaaaceaeg tteeaaatea gegaeatete teaaattgag 120
ateatagget ttttgaagat tgeteaaatt atgettetea tattgeatga geattttgaa 180
geeegegtea teaaceaaag eatttttee acceateaea atgattttat eattttettt 240
aaaatt

<210> 12

<211> 210

<212> PRT

<213> Caenorhabditis elegans

<400> 12

Met Lys Val Asn Val Ser Cys Ser Ser Val Gln Thr Thr Ile Asp Ile
1 5 10 15

Leu Glu Glu Asn Gln Gly Glu Asp Glu Ser Ile Leu Thr Leu Gly Gln
20 25 30

Leu Arg Asp Arg Ile Ala Thr Asp Asn Asp Val Asp Val Glu Thr Met
35 40 45

Lys Leu Leu His Arg Gly Lys Phe Leu Gln Gly Ala Asp Asp Val Ser 50 55 60

Leu Ser Thr Leu Asn Phe Lys Glu Asn Asp Lys Ile Ile Val Met Gly
65 70 75 80

Gly Lys Asn Ala Leu Val Asp Asp Ala Gly Phe Lys Met Leu Met Gln 85 90 95

Tyr Glu Lys His Asn Leu Ser Asn Leu Gln Lys Ala Tyr Asp Leu Asn 100 105 110

Leu Arg Asp Val Ala Asp Leu Glu Arg Gly Phe Leu Glu Lys Pro Lys
115 120 125

Gln Val Glu Met Gly Lys Lys Leu Glu Lys Lys Val Lys Tyr Phe Asn 130 135 140

Glu Glu Ala Glu Arg His Leu Glu Thr Leu Asp Gly Met Asn Ile Ile 145 150 155 160

Thr Glu Thr Thr Pro Glu Asn Gln Ala Lys Arg Asn Arg Glu Lys Arg 165 170 175

Lys Thr Leu Val Asn Gly Ile Gln Thr Leu Leu Asn Gln Asn Asp Ala 180 185 190

Leu Leu Arg Arg Leu Gln Glu Tyr Gln Ser Val Leu Asn Gly Asp Ile 195 200 205

Pro Glu 210

<210> 13

<211> 1377

<212> DNA

<213> Caenorhabditis elegans

<220>

<221> CDS

<222> (1)..(1377)

<400> 13

atg cca gtc gtg aac ata cca atc aaa ata ctt ggt cag aat caa tca 48
Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser
1 5 10 15

cat agt cga agt aac tcc tcg tct tct gtt gac aac gat cga aat caa 96
His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
20 25 30

caa tac cag cag gct cca aac gtg aat acc aat atg cat cat tcc aac 192
Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn
50 55 60

gga ttc tca cct aac ttc cca tct cgt agt cct att ccg gac ttt ccc 240 Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro 65 70 75 80

agt ttt tca tct ggg ttc cca aac gat tct gaa tgg tct tcg aat ttc 288 Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe 85 90 95

									aat Asn			336
	-								gga Gly			384
	-	_							cca Pro 140			432
									cag Gln			480
									caa Gln			528
		_							cca Pro			576
									gca Ala			624
									aaa Lys 220			672
									att Ile			720
								Glu	caa Gln			768
_				Leu			Thr		gga Gly			816
	_		Leu			Ser			gat Asp	Asn		864

													ttg Leu	912
													cgt Arg	960
		-											gaa Glu 335	1008
_	_												aga Arg	1056
_		_											cag Gln	1104
	-	-	_										gaa Glu	1152
													tac Tyr	1200
													aac Asn 415	1248
													atc Ile	1296
			Glu					Gln				Glu	aga Arg	1344
	-	Asp	_	_	gat Asp	-	Gln		gaa Glu	tag				1377

<210> 14

<211> 458

<212> PRT

<213> Caenorhabditis elegans

<400> 14

BNSDOCID <WC _0014106A1 >

Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser 1 5 10 15

His Ser Arg Ser Asn Ser Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
20 25 30

Pro Pro Gin Gin Pro Pro Gin Pro Gin Gin Gin Gin Ser Gin Gin 35

Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn 50 55 60

Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro 65 70 75 80

Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe 85 90 95

Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn 100 105 110

Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn 115 120 125

Pro Pro Met Gln Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gln 130 135 140

Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gln Ala Pro 145 150 155 160

Pro Gln Tyr Ser Gln Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr 165 170 175

Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg
180 185 190

Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu 195 200 205

Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp 210 215 220

Ser Gly Lys Gly Asp Glu Asn Leu Glu Glu Asn Ile Ala Lys Ile Thr 225 230 235 240

Ile Giy Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly
245 250 255

- Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg 260 265 270
- Gly Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys 275 280 285
- Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg 290 295 300
- Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser 305 310 315 320
- Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys 325 330 335
- Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr 340 345 350
- Lys Thr Val Gln Val Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys 355 360 . 365
- Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met 370 375 380
- Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met 385 390 395 400
- Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe 405 410 415
- Leu Lys Ile Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
 420 425 430
- Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr
 435 440 445
- Lys Ala Asp Leu Met Asp Asp Gln Ser Glu
 450 455
- <210> 15
- <211> 588
- <212> DNA
- <213> Schizosaccharomyces pombe

<220>

<221> CDS

<222→ (1)..(588)

<400> 15

atg tca gaa aag act agc aca gtt aca ata cac tat gga aat cag cga 48
Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg
1 5 10 15

ttt ccg gta gca gtc aat cta aat gag acg tta agt gaa ctg att gat 96
Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp
20 25 30

gat tta ctt gaa acg act gag att tct gag aag aaa gtc aag ctt ttt 144 Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe 35 40 45

tac gct ggc aag cgt tta aaa gac aaa aaa gcc tcg tta tca aaa ttg 192
Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu
50 55 60

ggt tta aaa aat cat agt aaa att cta tgt ata aga cca cat aag caa 240 Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln 65 70 75 80

caa cga ggt tcc aag gaa aaa gac acg gtt gag ccc gct ccg aaa gcg 288
Gin Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala
85 90 95

gaa gcg gag aat cct gta ttt tcg cgt att tct gga gaa ata aaa gcc 336 Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala 100 105 110

atc gat cag tat gtt gac aaa gaa ctt tcc ccc atg tac gac aat tac 384

Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr

115 120 125

gta aat aaa ccg tcg aac gat cca aag cag aaa aac aaa cag aaa cta 432 Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu 130 135 140

atg ata agt gaa cta ctt tta caa cag ctt tta aaa ttg gat gga gtt 480 Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val 145 150 155 160

gac gta ctg ggc agc gag aaa ttg cgt ttt gaa cgg aag caa ctt gtt 528 Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val 165 170

tct aag atc caa aaa atg ttg gat cac gtt gac caa aca age caa gaa 576 Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu 180 185 190

gtg gcc gca tag 588

Val Ala Ala 195

<210> 16

<211> 195

<212> PRT

<213> Schizosaccharomyces pombe

<400> 16

Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg
1 5 10 15

Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp 20 25 30

Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe 35 40 45

Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu 50 55 60

Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln 65 70 75 80

Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala 85 90 95

Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala 100 105 110

Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr
115 120 125

Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu 130 135 140

Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val 145 150 155 160

Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val 165 170 175

Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu 180 185 190

Val Ala Ala 195

<210> 17

<211> 621

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1)..(621)

<400> 17

atg tot ttt ttt acc cag ttg tgt tot atg gat aaa aaa tat tgg atc 48
Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile
1 5 10 15

tct cta gct gta ttg tca gtt act gtt ttg att agc gca tta ttg aaa 96
Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys
20 25 30

aag aga get act gaa ace gaa gat att gte gtt gtt cat tae gat gge 144
Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val Val His Tyr Asp Gly
35 40 45

gaa aag ttg aat ttt gtg ttg cga caa cca agg ctg aat atg gtt tct 192 Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser 50 55 60

tac act agt ttt ctt cgt cgc gtg tgc aac gca ttt tca gta atg ccc 240

Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro

65 70 75 80

gac aaa gcg tct ctc aag tta aac ggg gtg acc ctc aag gat ggt tca 288
Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser
85 90 95

ctt tcc gac caa aat gtg caa aat gga agt gaa tta gag ctc gaa tta 336 Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu Glu Leu 100 105 110

ccc aaa ctg agc ccg gca atg caa caa att gaa gca tat ata gat gag 384 Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu

	115				120			125		
ctt caa Leu Gln 130										432
ccc gct Pro Ala 145		Gln								480
aca ttg Thr Leu	-									528
gac cca Asp Pro	-	Arg								576
caa tat Gln Tyr									tga	621
<210> 18 <211> 20 <212> P! <213> So	06 RT chizosad	ccharc	omyce	es p	ombe					

<400> 18

Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile 1 5 10 15

Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys 20 25 30

Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val His Tyr Asp Gly
35 40 45

Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser 50 55 60

Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro 65 70 75 80

Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser 85 90 95

Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu

100 105 110

Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu 115 120 125

Leu Gln Gin Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser 130 135 140

Pro Ala Ser Ala Gln Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu 145 150 155 160

Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp 165 170 175

Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln 180 185 190

Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys 195 200 205

<210> 19

<211> 2534

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (307)..(2034)

<400> 19

geggagetee geatecaace eegggeegeg geeaacttet etggaetgga eeagaagttt 60

ctageoggee agttgctace teeetttate teeteettee cetetggeag egaggagget 120

atttccagac acttccaccc ctctctggcc acgtcacccc cgcctttaat tcataaaggt 180

georggegee ggetteeegg acaegtegge ggeggagagg ggeceaegge ggeggeeegg 240

ccagagacte ggegeeegga gecagegee egeaeeegeg eeceageggg cagaeeecaa 300

cccagc atg age gee gee ace cae teg eee atg atg eag gtg geg tee 348 Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser

1 5 10

ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg gag atc aag atc 396 Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile 15 20 25 30

-	_								cac			444
									ggc Gly			492
									tct Ser			540
_									cga Arg 90			588
						-	 _	-	cgg Arg			6 36
									ttc Phe			684
									ctg Leu			732
									gtg Val			780
-	-	_			-				 cgg Arg 170			828
_	-		-	_					gcc Ala			876
									ccg Pro			924
		-		Ile					cgg Arg			972

					-	_	-	acg Thr					-		, ,	1020
		_			-		-	tac Tyr		-		,		_	_	1068
					-			gca Ala		-					-	1116
-		_	-	-		-		tca Ser		-		-	_	_		1164
								gtg Val 295					-			1212
-	-					_	_	act Thr	_		-		-		-	1260
			-	-	_			cca Pro	_			-				1308
					Gln	gtg Val		cgc Ara				-				1356
					340			9	Буб	345				-,-	350	
-		_	_		cca			tct Ser	gag	345 aag	gta	gag	gtg	aaa	gtt	1404
Val	Ser	Gln	Lys	Pro 355 gtt	cca Pro	Pro tgt	Pro	tct	gag Glu 360	345 aag Lys agc	gta Val cct	gag Glu ggc	gtg Val	aaa Lys 365	gtt Val gct	1404
Val ccc Pro	Ser cct Pro	Gln gct Ala	Lys cca Pro 370	Pro 355 gtt Val	cca Pro cct Pro	Pro tgt Cys	Pro cct Pro	tct Ser cct Pro	gag Glu 360 ccc Pro	aag Lys agc ser	gta Val Cct Pro	gag Glu ggc Gly	gtg Val cct Pro 380	aaa Lys 365 tct Ser	gtt Val gct Ala	

_						•		_		-		_		ctg Leu		1596
415					420					425					430	
														aag		1644
Lys	val	GIN	GIĀ	435	GIU	GIN	Ala	vai	440	ASN	rne	GIU	GIY	Lys 445	Lys	
	_		,		_	-		_	-		-			gag		1692
Thr	Asp	Lys	450	ıyr	Leu	мес	11e	455	GIU	lyr	Leu	Int	460	Glu	Leu	
_	-	-	_			-				-	-	-		cgt		1740
Leu	мта	465	Asp	ser	Val	ASP	470	Giu	GIY	Arg	MIA	475	AGI	Arg	GIN	
-		-	-		_		-	_	_			-	_	aaa Lys		1788
MIG	480	Arg	ASP	Gly	vai	485	Гуз	Vai	GIII	1112	490	Беа	GI (d	пуз	Беа	
_	_		~		_	_				-	_	-		gaa Glu		1836
495	GIII	БуЗ	nia	110	500	Vai	110	Oly	0111	505	0111	V 3.1	. 7.	Old	510	
-		-			-	-	-	-		_	-	-		atg Met		1884
				515					520					525		
_		-		_	-	-	-		_			_		aat Asn	•	1932
	1		530				_, _	535	~,~	-,-			540			
-	-				-		-	_		-	_		-	gca Ala		1980
		545					550					555				
				_	_									gca Ala	gca Ala	2028
****	560	nan		261	Der	565	4114	nsp		110	570	11011		,, <u>.</u> .a	₹ 3 ± CI	
ccg Pro 575		cct	ctgc	cct	gtaa	aaat	ca g	actc	ggaa	c c g.	a tgt:	gtgc	ttt	aggga	aat	2084
ttt	aagt	tgc	atgc	attt	ca g	agac	ttta.	a gt	cagt	tggt	ttt	tatt	agc	tgct	tggtat	2144
gca	gtaa	ctt	gggt	ggag	gc a	aaac	acta	a ta	aaag	ggct	aaa	aagg.	aaa	atga	tgcttt	2204

cogttgottg ttotgoagoo otgtotactt gggcaccoo accacctgtt agotgtggtt 2324
gtgcactgtc ttttgtagot otggactgga ggggtagatg gggagtcaat tacccatcac 2384
ataaatatga aacatttato agaaatgttg coattttaat gagatgattt tottcatoto 2444
ataattaaaa tacctgactt tagagagagt aaaatgtgcc aggagccata ggaatatotg 2504
tatgttggat gactttaatg ctacatttto 2534

<210> 20

<211> 575

<212> PRT

<213> Homo sapiens

<400> 20

Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser Gly Asn
1 5 10 15

Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile Asp Pro
20 25 30

Gln Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr Thr Thr
35 40 45

Trp Asn Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr Pro Ser 50 55 60

Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro Ala Arg 65 70 75 80

Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile Pro Ile 85 90 95

Pro Val Leu His Glu Gly Ala Glu Asn Arg Gln Val His Pro Phe His
100 105 110

Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Glu Ala Ala 115 120 125

Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro Glu Thr 130 135 140

Thr Gln Pro Asp Lys Gln Cys Gly Gln Val Ala Ala Ala Ala Ala Ala

145					150					155					160
Gln	Pro	Pro	Ala	Ser 165	His	Gjy	Pro	Glu	Arg 170	Ser	Gln	Ser	Pro	Ala 175	Ala
Ser	Asp	Cys	Ser 180	Ser	Ser	Ser	Ser	Ser 185	Ala	Ser	Leu	Pro	Ser 190	Ser	Gly
Arg	Ser	Ser 195	Leu	Gly	Ser	His	Gln 200	Leu	Pro	Arg	Gly	Tyr 205	Ile	Ser	Ile
Pro	Val 210	Ile	His	Glu	Gln	Asn 215	Val	Thr	Arg	Pro	Ala 220	Ala	Gln	Pro	Ser
Phe 225	His	Lys	Ala	Gln	Lys 230	Thr	His	Tyr	Pro	Ala 235	Gln	Arg	Gly	Glu	Tyr 240
Gln	Thr	His	Gln	Pro 245	Val	Tyr	His	Lys	Ile 250	Gln	Gly	Asp	Asp	Trp 255	Glu
Pro	Arg	Pro	Leu 260	Arg	Ala	Ala	Ser	Pro 265	Phe	Arg	Ser	Ser	Val 270	Gln	Gly
		275					280				Ser	285			
Ser	Pro 290	Ser	Pro	Ile	Arg	Val 295	His	Thr	Val	Val	Asp 300	Arg	Pro	Gln	Gln
305					310					315	Gln				320
			-	325	-				330		Leu			335	
			340					345			Ser		350		
		35 5					360				Val	365			
Ala	Pro 370	Val	Pro	Cys	Pro	Pro 375		Ser	Pro	Gly	Pro 380	Ser	Ala	Val	Pro
385			_		390					395	Ala				400
Ala	Pro	Ala	Glu	Ala	Thr	Pro	Pro	Lys	Pro	Gly	Glu	Ala	Glu	Ala	Pro

405 410 415

Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val
420 425 430

Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp 435 440 445

Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala 450 460

Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg 465 470 475 480

Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln
485 490 495

Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro 500 505 510

Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly 515 520 525

Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp 530 535 540

Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser 545 550 555 560

Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala Pro 565 570 575

<210> 21

<211> 1966

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (43)..(1416)

<400> 21

BNSDOCID <WC 0014106A1 :>

cggtgggagc ggggcgggaa gcgcttcagg gcagcggatc cc atg tcg gcc ctg 500 Met Ser Ala Leu

agg cgc tcg ggc tac ggc ccc agt gac ggt ccg tcc tac ggc cgc tac 102

Arg 5	Arg	Ser	Gly	Tyr	Gly 10	Pro	Ser	Asp	Gly	Pro 15	Ser	Tyr	Gly	Arg	Tyr 20	
											cca Pro					150
											att Ile					198
-											gga Gly					246
	-										cca Pro 80					294
_											cct Pro					342
											cct Pro					390
				Pro							ttg Leu					438
			Tyr					Pro			cct Pro		Ala			486
_		Tyr					Tyr				tat Tyr 160					534
	Ser					Ser					tct Ser					582
					Trp					Glr	gac Asp				Glu	630
gca	a cc	c cct	ctt	agg	g ggç	g caç	g gtt	cca	gga	tat	ccç	cct	tca	cag	aac	6 78

Ala	Pro	Pro	Leu 200	Arg	Gly	Gln	Val	Pro 205	Gly	Tyr	Pro	Pro	Ser 210	Gln	Asn	
					ccc Pro											726
					ccg Pro											774
		-			gga Gly 250											822
					ccc Pro											870
					ggc											918
					tct Ser											966
					ttt Phe											1014
															cag Gln 340	1062
					cag Gln											1110
				Gln	agt Ser											1158
			Thr		ccg Pro			Lys					Val			1206
aag	gtc	cag	tat	ctt	gaa	caa	gaa	gta	gaa	gaa	ttt	gta	gga	aaa	aag	1254

Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe Val Gly Lys Lys

425

gcc aga aaa gag gct gtt tgt aag att cag gcc ata ctg gaa aaa tta 1398 Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu Lys Leu 440 445 450

430

435

gaa aaa aaa gga tta tga aaggatttag aacaaagtgg aagcctgtta 1446 Glu Lys Lys Gly Leu 455

<210> 22

<211> 457

<212> PRT

<213> Homo sapiens

<400> 22

Met Ser Ala Leu Arg Arg Ser Gly Tyr Gly Pro Ser Asp Gly Pro Ser 1 5 10 15

Tyr Gly Arg Tyr Tyr Gly Pro Gly Gly Gly Asp Val Pro Val His Pro

20 25 30

Pro Pro Pro Leu Tyr Pro Leu Arg Pro Glu Pro Pro Gln Pro Pro Ile 35 40 45

Ser Trp Arg Val Arg Gly Gly Gly Pro Ala Glu Thr Thr Trp Leu Gly 50 55 60

Glu Gly Gly Gly Gly Asp Gly Tyr Tyr Pro Ser Gly Gly Ala Trp Pro 65 70 75 80

Glu Pro Gly Arg Ala Gly Gly Ser His Gln Glu Gln Pro Pro Tyr Pro 85 90 95

Ser Tyr Asn Ser Asn Tyr Trp Asn Ser Thr Ala Arg Ser Arg Ala Pro 100 105 110

Tyr Pro Ser Thr Tyr Pro Val Arg Pro Glu Leu Gln Gly Gln Ser Leu 115 120 125

Asn Ser Tyr Thr Asn Gly Ala Tyr Gly Pro Thr Tyr Pro Pro Gly Pro 130 135 140

Gly Ala Asn Thr Ala Ser Tyr Ser Gly Ala Tyr Tyr Ala Pro Gly Tyr 145 150 155 160

Thr Gln Thr Ser Tyr Ser Thr Glu Val Pro Ser Thr Tyr Arg Ser Ser 165 170 175

Gly Asn Ser Pro Thr Pro Val Ser Arg Trp Ile Tyr Pro Gln Gln Asp 180 185 190

Cys Gln Thr Glu Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro 195 200 205

Pro Ser Gln Asn Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp 210 215 220

Gly Asn Arg Ser Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu 225 230 235 240

Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly Arg Tyr Pro 245 250 255

Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr 260 265 270

Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro

275 280 285

Ser Pro Pro Val Gin Gin Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln 290 295 300

Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln 305 310 315 320

Tyr Glu Ser Ser Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp 325 330 335

Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr 340 345 350

Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu Pro Glu Glu 355 360 365

Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile 370 375 380

His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe 385 390 395 400

Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu 405 410 415

Thr Lys Glu Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp
420 425 430

Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile 435 440 445

Leu Glu Lys Leu Glu Lys Lys Gly Leu 450 455

<210> 23

<211> 4308

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (247)..(1590)

<400> 23

coccoccc coccccc congaagacg cocggagegg etgetgeage cagtagegge 60

contrology getgeocoge teagacetag tegggagggg tgegaggeat geagetgggg 12	O
geocagetee ggtgeegeae eeegtaaagg getgatette caeetegeea eeteageeae 18	Ю
gggacgccaa gaccgcatcc aattcagact tettttggtg ettgtgaaac tgaacacaac 24	0
aaaagt atg gat atg gga aac caa cat cct tct att agt agg ctt cag 28 Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln 1 5 10	8
gaa atc caa aag gaa gta aaa agt gta gaa cag caa gtt atc ggc ttc 33 Glu Ile Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe 15 20 25 30	16
agt ggt ctg tca gat gac aag aat tac aag aaa ctg gag agg att cta 38 Ser Gly Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu 35 40 45	4
aca aaa cag ctt ttt gaa ata gac tct gta gat act gaa gga aaa gga 43 Thr Lys Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly 50 55 60	12
gat att cag caa gct agg aag cgg gca gca cag gag aca gaa cgt ctt 48 Asp Ile Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu 65 70 75	10
ctc aaa gag ttg gag cag aat gca aac cac cca cac cgg att gaa ata 52 Leu Lys Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile 80 85 90	:8
cag aac att ttt gag gaa gcc cag tcc ctc gtg aga gag aaa att gtg 57 Gln Asn Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val 95 100 105 110	16
cca ttt tat aat gga ggc aac tgc gta act gat gag ttt gaa gaa ggc 62 Pro Phe Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly 115 120 125	14
atc caa gat atc att ctg agg ctg aca cat gtt aaa act gga gga aaa 67 Ile Gln Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys 130 135 140	12
atc tcc ttg cgg aaa gca agg tat cac act tta acc aaa atc tgt gcg 72 Ile Ser Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala 145 150 155	20
gtg caa gag ata atc gaa gac tgc atg aaa aag cag cct tcc ctg ccg 76 Val Gln Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro	58

	160					165				170				
										atc Ile				816
										gca Ala				864
-										tgt Cys				912
_		_								ggc Gly				960
_					-	-	_	-	_	aac Asn 250				1008
	_	_	_	_		_	_	-		aa a Lys	_		_	1056
-										gtc Val				1104
										aac Asn				1152
	_	-				-	-			att Ile				1200
										gaa Glu 330				1248
-	-									att Ile				1296
-				-	_	-		-	_	gag Glu				1344

WO 00/14106	PCT/US99/21053

355	360	365
aaa gcc gtc tgg aac gtc c Lys Ala Val Trp Asn Val L 370		
gtt ctt tca ttt gat gga a Val Leu Ser Phe Asp Gly A 385		
gaa gag ctg ctc acc aag c Glu Glu Leu Leu Thr Lys G 400	Sln Leu Leu Ala Leu A	
cag gga gaa gag aag tgt a Gln Gly Glu Glu Lys Cys I 415 420		
gcg cag aat att ctc agc t Ala Gln Asn Ile Leu Ser 1 435		
tac tga aataccagag atctca Tyr	acttt tgatactgtt ttgd	cacttca tatgtgcttc 1640
tatgtataga gagettteag tte	cattgatt tatacgtgca t	atttcagtc tcagtattta 1700
tgattgaagc aaattctatt cac	gtatotgo tgottttgat ç	gttgcaagac aaatatcatt 1760
acagcacgtt aacttttcca tto	cggatcat tatctgtatg a	atgtggtgtg gtttgtttgg 1820
tttgtccttt tttttgcgtt tt	taatcaga aaacaaaata q	gaggcagctt ttgtagattt 1880
taaatgggtt gtgcaagcat ta	aaatgcag gtctttcaga a	atctagaact aggcataacc 1940
ttacataata ctaggaaaat ta	tgagaaag gggaaatttt t	tggttaaata agagtaaggt 2000
tcaaacacaa gcagtacatg tt	ctgtttca ttatgctcga 1	tagaaggett ttttttcact 2060
tataaggeet gattggteet ac	ccagetta acggggtggg	gtttttttgt ttgttcagac 2120
agtotgttot tttgtaaaca tt	tttagttg gaaaaacagc .	atotgoattt tooccateet 2180
ctacgtttta gagaggaatc tt	gtttttgt gtgcaacata	agaaaattat gaaaactaat 2240
agccaaaaaa cctttgagat tg	cattaaag agaagggata	aaggaccagc aataatacct 2300

BNSDOCID <WC _ 0014106A1 . >

aaggatggga acatttaaat taagttaatg ggcctttaaa aaaaaaaaag gaaacactca 2420 tacctgtagt tggaggatga atactggaga ogggttacca atgtcaggtt atactaaaac 2460taaatcagaa agtetgaatg tageacataa tggttetett etgttgteea aggetgtaaa 2540 atggacaged ttgtcacace teecoggtge tgttttacaa egtgagggta gaegetgtea 2600 gtaacccaga gggaccagge ettectaggt tttetaggea gteagetgtt aaccaeteae 2660 ttagtaaatg teataactae aeetgeteea ggaeeaatea gtgaaaeetg eteggaatta 2720 aaggetteet etgggtgeet getgaacaae tgageteatg teatgggeat gtggtggttt 2780 ctctqttqcc tqaaagagcc attaaagtca gtcgtgcgtg aagcatctct cttctaaagg 2840 atqtqtattt ccataaatgc tttctgagga tccggtacaa aatgatttcc caaagttctg 2900 aagtgeettg agaacatgtg ggteegagtg ttataacaga etecteece gggteacett 2960 ttgcctggtc atcctgttag agtacatett tggaaateca gggtaatatt etettteaga 3020 gatgctcatt gtgtaactct gtgtagggag atagtcactt taaacagctc aaagtagcta 3080 gctaaaggag tagccttaaa tacctaaaag atgacagaag catagccctt aacaaatctt 3140 cagettgtet eteagtattt eccaateatg aaaateeett getatgtett teetaetaga 3200 aatgttetag aategetgga eggtggggte agagggeagt eggtatttag geegtgaget 3260 teceatacta etgeaggtee aacteetgge aacegeggge teaaggeagg teattggaat 3320 ccacgttttg gccacagtag ttgtaggatt gcttttctgt atcataattt tagaatgctc 3380 ttaaaatctt gaggaagagt ttttattttt tatttatttt tgagatggag tctctgttgc 3440 ccaggctgca gtgcagtggt gccatctcag ctcactgcaa cctccacctc ccaggttcaa 3500 gcgattctcc tgcctcagcc acctgagtag ctgggagtac aggcatgtgg caccatgcct 3560 ggctaatttt tgtattttta atagagttga gatttcacca tgatggtcag gctggtctcg 3620 aactootgae etegtgatee geoegeeteg geoeceeaaa gtgetgggat taaegggtgt 3680 gagecaegge geceagecea ggaagagttt ttaaattaga getetgttta attataceae 3740 tgggaaatca tggttacgct tcaggcatat tcttccccag agtactactt acattttaaa 3800

tttcattttg taaagttaaa tgtcagcatt ccctttaaaa gtgtccattg ttcttgaaa 3860 gtagacgttt cagtcattct tttcaaacaa gtgtttgtg accttttgcc aagctgtggg 3920 catcgtgtg gagtacaggg tgctcagctc ttccaccgtc attttgaatt gttcacatgg 3980 gtaattggtc atggaaatga tcagattgac cttgattgac tgtcaggcat ggctttgttt 4040 ctagtttcaa tctgttctcg ttccttgtac cggattattc tactcctgca atgaaccctg 4100 ttgacaccgg atttagctct tgtcggcctt cgtggggagc tgtttgtgt aatatgagc 4160 actgcatga attctaaac tgggcttgtc acattgtatt gtattttgt gatctgtaat 4220 gaaaaagaatc tgtactgcaa gtaaaaccta ctccccaaaa atgtgtggct ttgggtctgc 4280 attaaacgct gtagtccatg ttcatgcc

<210> 24

<211> 447

<212> PRT

<213> Homo sapiens

<400> 24

Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln Glu Ile 1 5 10 15

Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe Ser Gly
20 25 30

Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu Thr Lys
35 40 45

Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly Asp Ile
50 55 60

Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu Leu Lys
65 70 75 80

Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile Gln Asn 85 90 95

Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val Pro Phe 100 105 110

Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly Ile Gln

		115					120					125			
Asp	Ile 130	Ile	Leu	Arg	Leu	Thr 135	His	Val	Lys	Thr	Gly 140	Gly	Lys	Ile	Ser
Leu 145	Arg	Lys	Ala	Arg	Tyr 150	His	Thr	Leu	Thr	Lys 155	Ile	Cys	Ala	Val	Gln 160
Glu	Ile	Ile	Glu	Asp 165	Cys	Met	Lys	Lys	Gln 170	Pro	Ser	Leu	Pro	Leu 175	Ser
Glu	Asp	Ala	His 180	Pro	Ser	Val	Ala	Lys 185	Ile	Asn	Phe	Val	Met 190	Cys	Glu
Val	Asn	Lys 195	Ala	Arg	Gly	Val	Leu 200	Ile	Ala	Leu	Leu	Met 205	Gly	Val	Asn
Asn	Asn 210	Glu	Thr	Cys	Arg	His 215	Leu	Ser	Cys	Val	Leu 220	Ser	Gly	Leu	Ile
Ala 225	Asp	Leu	Asp	Ala	Leu 230	Asp	Val	Cys	Gly	Arg 235	Thr	Glu	Ile	Arg	Asn 240
Tyr	Arg	Arg	Glu	Val 245	Val	Glu	Asp	Ile	Asn 250	Lys	Leu	Leu	Lys	Tyr 255	Leu
Asp	Leu	Glu	Glu 260	Glu	Ala	Asp	Thr	Thr 265	Lys	Ala	Phe	Asp	Leu 270	Arg	Gln
Asn	His	Ser 275	Ile	Leu	Lys	Ile	Glu 280	Lys	Val	Leu	Lys	Arg 285	Met	Arg	Glu
Ile	Lys 290	Asn	Glu	Leu	Leu	Gln 295	Ala	Gln	Asn	Pro	Ser 300	Glu	Leu	Tyr	Leu
Ser 305	Ser	Lys	Thr	Glu	Leu 310	Gln	Gly	Leu	Ile	Gly 315	Gln	Leu	Asp	Glu	Val 320
Ser	Leu	Glu	Lys	Asn 325	Pro	Cys	Ile	Arg	Glu 330		Arg	Arg	Arg	Ala 335	Val
Ile	Glu	Val	Gln 340		Leu	Ile	Thr	Tyr 345		Asp	Leu	Lys	Glu 350	Ala	Leu
Glu	Lys	Arg 355	_	Leu	Phe	Ala	Cys 360		Glu	His	Pro	Ser 365		Lys	Ala

Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val Leu

 Ser
 Phe
 Asp
 Gly
 Asn
 Arg
 Thr
 Asp
 Lys
 Asn
 Tyr
 Ile
 Arg
 Leu
 Glu
 Glu
 Glu
 Glu
 Asp
 Ala
 Leu
 Ala
 Leu
 Asp
 Ala
 Val
 Asp
 Pro
 Gln
 Gly
 Gly
 Ala
 Ala
 Arg
 Lys
 Gln
 Ala
 Val
 Arg
 Leu
 Ala
 Gln
 Ala
 Ala
 Gln
 Ala
 Arg
 Leu
 Ala
 Arg
 Lys
 Gln
 Ala
 Arg
 Leu
 Ala
 Arg
 Leu
 Leu
 Ala
 Arg
 Leu
 Leu
 Ala
 Arg
 Leu
 Ala
 Arg
 Leu
 Arg
 Leu
 Arg
 Eu
 Leu
 Arg
 Leu
 Arg
 Eu
 Eu
 Arg
 Eu
 Arg
 Eu
 Arg
 Eu
 Eu
 Arg
 Eu
 Arg
 Eu
 Arg
 Eu
 Eu

INTERNATIONAL SEARCH REPORT

International application No. PCT/US99/21053

A. CLASSIFICATION OF SUBJECT MATTER						
IPC(6) :07N 21/02; C07K 1/00 US CL :530/387.1, 350; 435/6, 7/1; 536/23.1						
According to International Patent Classification (IPC) or to both national classification and IPC						
B. FIEL	DS SEARCHED					
Minimum de	ocumentation searched (classification system followed	by classification symbols)				
U.S. : :	530/387.1, 350; 435/6, 7/1; 536/23.1					
Documentat	ion searched other than minimum documentation to the	extent that such documents are included	in the fields searched			
Electronic d	ata base consulted during the international search (nar	ne of data base and, where practicable,	search terms used)			
C. DOC	UMENTS CONSIDERED TO BE RELEVANT					
Category*	Citation of document, with indication, where app	ropriate, of the relevant passages	Relevant to claim No.			
X	US 5,652,223 A (KOHN ET AL) 29 J document.	2-5, 14, 32-34				
X	Database Genbank-EST, National C Accession No. AA693697, HILLIER human EST Project, 16 December 199	2				
X	Database Genbank-EST, National Center for Biotech. Info., Accession No. AA456862, NCI_CGAP, 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index,' 15 August 1997, see entire reference.					
X Further documents are listed in the continuation of Box C. See patent family annex.						
Special categories of cited documents:						
	becoment defining the general state of the art which is not considered be of particular relevance	the principle or theory underlying th				
E ca	urlier document published on or after the international filing date	"X" document of particular relevance, the considered novel or cannot be considered.	ne claimed invention cannot be ered to involve an inventive step			
	ocument which may throw doubts on priority claim(s) or which is ted to establish the publication date of another citation or other	when the document is taken slone				
	pecial reason (as specified)	 'Y' document of particular relevance; the considered to involve an inventor 	e step when the document is			
	ocument referring to an oral disclosure, use, exhibition or other leans	combined with one or more other su- being obvious to a person skilled in	one or more other such documents, such combination to a person skilled in the art			
	ocument published prior to the international filing date but later than	"&" document member of the same pater	nt family			
Date of the actual completion of the international search Date of mailing of the international search report						
24 NOVEMBER 1999 19 JAN 2000						
Commissi Box PCT	mailing address of the ISA/US oner of Patents and Trademarks on, D.C. 20231	SHEELA J. HUFF				
Facsimile No. (703) 305-3230		Telephone No. (703) 308-0196				

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No PCT/US99/21053

Box 1 Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)				
This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:				
1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:				
2. X Claims Nos.: 1, 13, 24, 25 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: No.meaningful search could be carried out because no limitations could be placed on the sequence				
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).				
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)				
This International Searching Authority found multiple inventions in this international application, as follows:				
1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.				
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.				
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:				
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:				
Remark on Protest The additional search fees were accompanied by the applicant's protest No protest accompanied the payment of additional search fees				

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No PCT/US99/21053

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	
X	Database Genbank, National Center for Biotech. Info., Accession No. G29287, MYERS, R.M., 04 October 1996, see entire reference.		
X	Database Genbank, National Center for Biotech. Info., Accession No. G06974, HUDSON, T., "Whitehead Institute.MIT Center for Genome Research,'19 October 1995, see entire reference.	2,4	
X	Database Genseq, Derwent, Alexandria, Virginia, Accession No. V81267, OTSUKA PHARM CO LTD, 'New Bcl-2 interaction prrtein gene (Bis)- useful for elucidation of the molecular mechanism of apoptosis, and in diagnosis, prevention and treatment of diseases,' 15 December 1998 see enire reference.	2-5	
X	Database, Geneseq, Derwent, Alexandria, Virginia, Accession No. T19051, MATSUBARA ET AL., "Identifying gene signatures in 3'-directed human cDNA library,' 01 June 1995, see entire reference.	2,4	
X	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. Q90296, LA JOLLA CANCER RES FOUN. 'Human Bcl-2-associated protein BAG-1 cDNA,'18 May 1995 see entire reference.	2-5,14	

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07K 1/00	AI	(11) International Publication Number:(43) International Publication Date:	WO 00/14106 16 March 2000 (16.03.00)	
(21) International Application Number: PCT/US (22) International Filing Date: 9 September 1999 ((81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).			
 (30) Priority Data: 09/150,489 9 September 1998 (09.09.98) (71) Applicant: THE BURNHAM INSTITUTE [US/US]; Torrey Pines Road, La Jolla, CA 92037 (US). (72) Inventors: REED, John, C.; 17044 El Camino Real Santa Fe, CA 92067 (US). TAKAYAMA, Shin Stratford Court #3, Del Mar, CA 92014 (US). (74) Agents: WONG, James, J. et al.; Campbell & Flores L 700, 4370 La Jolla Village Drive, San Diego, C (US). 	Before the expiration of the t claims and to be republished i. amendments.	Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.		
(54) Title: NOVEL BAG PROTEINS AND NUCLEIC A	.CID M	OLECULES ENCODING THEM		

(57) Abstract

The present invention provides a family of BAG-1 related proteins from humans (BAG-1L, BAG-1, BAG-2, BAG-3, BAG-4 and BAG-5), the invertebrate *C. elegans* (BAG-1, BAG-2) and the fission yeast *S. pombe* (BAG-1A, BAG-1B) and the nucleic acid molecules that encode them.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
\mathbf{BG}	Bulgaria	HU	Hungary	ML	Mali	77	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukrame
BR	Brazil	11.	Israel	MR	Mauritania	$\mathbf{U}\mathbf{G}$	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	$\mathbf{U}\mathbf{Z}$	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CC	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Canteroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
cu	Cuba	KZ	Kazakstan	RO	Romania		
CN	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE.	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Srr Lanka	SE	Sweden		
EE.	Estonia	LR	Liberia	SG	Singapore		

NOVEL BAG PROTEINS AND NUCLEIC ACID MOLECULES ENCODING THEM

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY-SPONSORED RESEARCH AND DEVELOPMENT

This invention was made with government support under grant number CA-67329 awarded by the National Institutes of Health. The United States Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention relates generally to the fields of molecular biology and molecular medicine and more specifically to a novel family of proteins that can regulate protein folding. The functions of these proteins are potentially diverse, including promoting tumor cell growth and metastasis.

BACKGROUND INFORMATION

The Hsc70/Hsp70-family of molecular chaperones participate in protein folding reactions, controlling protein bioactivity, degradation, complex assembly/disassembly, and translocation across membranes. These proteins interact with hydrophobic regions within target proteins via a carboxyl (C)-terminal peptide binding domain, with substrate binding and release being controlled by the N-terminal ATP-binding domain of Hsc70/Hsp70. Hsc70/Hsp70-assisted folding reactions are accomplished by repeated cycles of peptide binding, refolding, and release,

10

PCT/US99/21053 WO 00/14106

2

which are coupled to ATP hydrolysis by the ATP-binding domain (ATPase) of Hsc70/Hsp70 and by subsequent nucleotide exchange. The chaperone activity of mammalian Hsc70/Hsp70 is regulated by partner proteins that either modulate the 5 peptide binding cycle or that target the actions of these chaperones to specific proteins and subcellular compartments. DnaJ-family proteins (Hdj-1/Hsp40; Hdj-2; Hdj-3) stimulate the ATPase activity of Hsc70/Hsp70, resulting in the ADP-bound state which binds tightly to peptide substrates. The Hip protein collaborates with Hsc70/Hsp70 and DnaJ homologues in stimulating ATP hydrolysis, and thus also stabilize Hsc70/Hsp70 complexes with substrate polypeptides, whereas the Hop protein may provide co-chaperone functions through interactions with the C-terminal peptide binding domain.

The Bcl-2 associated athanogene-1 (bag-1) is named from the Greek word athanos, which refers to anti-cell death. BAG-1 was previously referred to as Bcl-2-associated protein-1 (BAP-1) in U.S. Patent No. 5,539,094 issued July 23, 1996, which is incorporated 20 herein by reference. In this earlier patent, BAG-1 is described as a portion of the human BAG-1 protein, absent the N-terminal amino acids 1 to 85. In addition, a human protein essentially identical to human BAG-1 was described 25 by Zeiner and Gehring, (Proc. Natl. Acad. Sci., USA 92:11465-11469 (1995)). Subsequent to the issuance of U.S. Patent 5,539,094 the N-terminal amino acid sequence from 1 to 85 of human BAG-1 was reported.

BAG-1 and its longer isoforms BAG-1M (Rap46) and 30 BAG-1L are recently described Hsc70/Hsp70-regulating proteins. BAG-1 competes with Hip for binding to the Hsc70/Hsp70 ATPase domain and promotes substrate release. BAG-1 also reportedly stimulates Hsc70-mediated ATP

10

15

3

hydrolysis by accelerating ADP/ATP exchange, analogous to the prokaryotic GrpE nucleotide exchange protein of the bacterial Hsc70 homologue, DnaK. Gene transfection studies indicate that BAG-1 proteins can influence a wide variety of cellular phenotypes through their interactions with Hsc70/Hsp70, including increasing resistance to apoptosis, promoting cell proliferation, enhancing tumor cell migration and metastasis, and altering transcriptional activity of steroid hormone receptors.

Despite the notable progress in the art, there remains an unmet need for the further identification and isolation of additional homologous BAG protein species, and the nucleic acid molecules and/or nucleotide sequences that encode them. Such species would provide additional means by which the identity and composition of the BAG domain, that is, the portion of the protein that is influencing or modulating protein folding, could be identified. In addition, such species would be useful for identifying agents that modulate apoptosis as candidates for therapeutic agents, in particular, anticancer agents. The present invention satisfies these need, as well as providing substantial related advantages.

SUMMARY OF THE INVENTION

The present invention provides a family of BAG-1 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO: 4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate C.elegans [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast S.pombe [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)] and the nucleic acid molecules that encode them.

4

Another aspect of the present invention provides an amino acid sequence present in the family of BAG-1 related proteins, that modulates Hsc70/Hsp70 chaperone activity, that is, the BAG domain.

Another aspect of the present invention provides novel polypeptide and nucleic acid compositions and methods useful in modulating Hsc70/Hsp70 chaperone activity.

Another aspect of the present invention is directed to methods for detecting agents that modulate the binding of the BAG family of proteins, such as BAG-1 (beginning at residue 116 of SEQ ID NO:2), and related proteins with the Hsc70/Hsp70 Family of proteins or with other proteins that may interact with the BAG-Family proteins.

Still another aspect of the present invention is directed to methods for detecting agents that induce the dissociation of a bound complex formed by the association of BAG-Family proteins with Hsc70/Hsp70 Family molecule chaperones or other proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the full length cDNA sequence for human BAG-1 (SEQ ID NO:1) protein with the corresponding amino acid sequence (SEQ ID NO:2). Within the full length sequence are included the overlapping sub-sequences of BAG-1 (beginning at nucleotide 391), BAG-1M [beginning at nucleotide 260 of (SEQ ID NO:2)], and BAG-1L [beginning at nucleotide 46 of (SEQ ID NO:2)].

20

5

Figures 2A and 2B combined shows the full length cDNA sequence (SEQ ID NO:3) aligned with the corresponding amino acid residues for human BAG-2 protein (SEQ ID NO:4).

Figure 3 shows a cDNA sequence (SEQ ID NO:5) saligned with the corresponding amino acid residues for human BAG-3 protein (SEQ ID NO:6).

Figure 4 shows the a cDNA sequence (SEQ ID NO:7) aligned with the corresponding amino acid residues for human BAG-4 protein (SEQ ID NO:8).

Figure 5 shows a cDNA sequence (SEQ ID NO:9) aligned with the corresponding amino acid residues for human BAG-5 protein (SEQ ID NO:10).

Figure 6A shows the full length cDNA sequence for \mathcal{C} . elegans BAG-1 protein (SEQ ID NO:11).

Figure 6B shows the 210 amino acid sequence for C. elegans BAG-1 protein (SEQ ID NO:12).

Figure 7A shows the full length cDNA sequence for \mathcal{C} . elegans BAG-2 protein (SEQ ID NO:13).

Figure 7B shows the 458 amino acid sequence for 20 C. elegans BAG-2 protein (SEQ ID NO:14).

Figure 8A shows the full length cDNA sequence for S. pombe BAG-1A protein (SEQ ID NO:15).

Figure 8B shows the 195 amino acid sequence for $S.\ pombe\ BAG-1A$ protein (SEQ ID NO:16).

6

Figure 9A shows the full length cDNA sequence for S. pombe BAG-1B protein (SEQ ID NO:17).

Figure 9B shows the 206 amino acid sequence for S. pombe BAG-1B protein (SEQ ID NO:18).

Figure 10 shows the topologies of the BAG-family 5 proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEO ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10); S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and C. elegans BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). (A) The relative 10 positions of the BAG domains are shown in black, ubiquitinlike regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like nuclear localization sequence are also shown. (B) The amino acid sequences of the BAG domain for human BAG-1 (SEQ ID NO:2), 15 BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10), S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18), and C. elegans BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) are aligned demonstrating their homology. Black and gray shading represent identical 20 and similar amino acids, respectively.

Figure 11 shows assays demonstrating the interaction of BAG-family proteins with Hsc70/ATPase. (A) Two-hybrid assays using yeast expressing the indicated fusion proteins. Blue color indicates a positive interaction, resulting in activation of the lacZ reporter gene. (B) In vitro protein assays using GST-fusion proteins and 35S-labeled in vitro translated proteins. (C) Co-immunoprecipitation assays using anti-Flag or IgG1 control antibodies and lysates from 293T cells expressing Flag-tagged BAG-1 (beginning at residue 116 of SEQ ID

25

7

NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), Daxx, or Apaf-1.

Figure 12 shows surface plasmon resonance analysis of EAG-family protein interactions with 5 Hsc70/ATPase. (A) SDS-PAGE analysis of purified recombinant proteins. (B) Representative SPR results of biosensor chips containing immobilized BAG proteins with and without maximally bound Hsc70/ATPase.

Figure 13 shows representative SPR results for biosensor chips containing immobilized BAG-1 (beginning at residue 116 at SEQ ID NO:2), BAG-1(ΔC), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) proteins. Hsc70/ATPase was flowed over the chips (arrow/left) until maximal binding was reached (response units), then flow was continued without Hsc70/ATPase (arrow/right). For BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6), Hsc70 was injected at 0.0175, 0.035, 0.07, 0.14, and 0.28 μM.

Figure 14 shows BAG-family protein modulation of Hsc70 chaperone activity. (A) Protein refolding assay of 20 chemically-denatured luciferase by Hsc70 plus DnaJ in the absence or presence of BAG and BAG-mutant proteins. Concentration-dependent inhibition of Hsc70-mediated protein refolding by BAG-family proteins [BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6)] but not by BAG-mutant (BAG-1 (Δ C). 25 Hsc70/Hsp40-mediated refolding of heat-denatured luciferase was assayed in the presence of (black bars) or absence of (striped bars) of 1.8 μM Hip, with (lanes 3-10) or without (lanes 1,2) various BAG-family proteins (1.8µM) indicated (mean ±SE; n=3). A control (CNTL) is shown (lane 30 1) in which Hsc70 was replaced with an equivalent amount of BSA.

8

Figure 15A shows an expanded cDNA sequence for human BAG-3 protein (SEQ ID NO:19).

Figure 15B shows the corresponding amino acid residues for the human BAG-3 protein (SEQ ID NO:20) of 5 Figure 15A.

Figure 15C shows the expanded cDNA sequence (SEQ ID NO:19) aligned with the corresponding amino acid residues for human BAG-3 protein of Figure 15A (SEQ ID NO:20).

10 Figure 16A shows an expanded cDNA sequence for human BAG-4 protein (SEQ ID NO:21).

Figure 16B shows the corresponding amino acid residues for the human BAG-4 protein of Figure 16A (SEQ ID NO:22).

Figure 16C shows the expanded cDNA sequence (SEQ ID NO:21) aligned with the corresponding amino acid residues for human BAG-4 protein of Figure 16A (SEQ ID NO:22).

Figure 17A shows an expanded cDNA sequence for 20 human BAG-5 protein (SEQ ID NO:23).

Figure 17B shows the corresponding amino acid residues for the human BAG-5 protein of Figure 17A (SEQ ID NO:24).

Figure 17C shows the expanded cDNA sequence (SEQ 25 ID NO:23) aligned with the corresponding amino acid residues for human BAG-5 protein of Figure 17A (SEQ ID NO:24).

9

Figure 18 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), expanded BAG-3 (SEQ ID NO:20), expanded BAG-4 (SEQ ID NO:22), expanded BAG-5 (SEQ ID NO:24); S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and C. elegans BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like nuclear localization sequence are also shown.

Definitions

10

The term "apoptosis", as used herein, refers to the process of programmed cell death, although not all programmed cell deaths occur through apoptosis, as used 15 herein, "apoptosis" and "programmed cell death" are used interchangeably.

The term "tumor cell proliferation", as used herein refers to the ability of tumor cells to grow and thus expand a tumor mass.

The term "cell migration", as used herein refers to the role cell motility plays in the invasion and potentially metastasis by tumor cells.

The term "metastasis", as used herein refers to the spread of a disease process from one part of the body to another, as in the appearance of neoplasms in parts of the body remote from the site of the primary tumor; results in dissemination of tumor cells by the lymphatics or blood vessels or by direct extension through serious cavitites or subarachnoid or other spaces.

10

The term "steroid hormone receptor function", as used herein refers to physiological, cellular and molecular functioning of receptors sites that bind with steroid hormones.

The term "substantially purified", as used herein, refers to nucleic acid or amino acid sequence that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Nucleic acid molecule" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single or double stranded, and represent the sense or antisense strand.

"Hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The terms "complementary" or "complementarity", 20 as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T binds to the complementary sequence "T-C-A".

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid and is referred to using the functional term "substantially homologous." The inhibition of

1.5

11

hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency.

The term "antisense", as used herein, refers to nucleotide sequences which are commplementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which 15 permits the synthesis of a complementary strand. introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant 20 phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense, and "positive" is sometimes used in reference to the sense strand.

"Amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited herein this term excludes an amino acid sequence of a naturally occurring protein. "Amino acid sequence", "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

PCT/US99/21053 **WO** 00/14106

12

The term "functional fragments" or "fragments", as used herein, with regard to a protein refers to portions of that protein that are capable of exhibiting or carrying out the activity exhibited by the protein as a whole. The 5 portions may range in size from three amino acid residues to the entire amino acid sequence minus one amino acid. For example, a protein "comprising at least a functional fragment of the amino acid sequence of SEQ ID NO:1", encompasses the full-length of the protein of SEQ ID NO:1 and portions thereof.

A "derivative" of a BAG protein, as used herein, refers to an amino acid sequence that is alterd by one or more amino acids. The derivative may have "conservative" changes, wherein a substituted amino acid has similar 15 structural or chemical properties, e.g., substitution of an apolar amino acid with another apolar amino acid (such as replacement of leucine with isoleucine). The derivative may also have "nonconservative" changes, wherein substituted amino acid has different but sufficiently similar structural or chemical properties that permits such 20 a substitution without adversely effecting the desired biological activity, e.g., replacement of an amino acid with an uncharged polar R group with an amino acid with an apolar R group (such as replacement of glycine with 25 tryptophan), or alternatively replacement of an amino acid with a charged R group with an amino acid with an uncharged Polar R group (such as replacement of lysine with asparagine).

10

15

13

Amino Acids - Apolar R Groups

Amino Acid	Radical	Abbrevi	lations
		3-Letter	1-Letter
alanine	methyl	ala	A
valine	2-propyl	aal	V
leucine	2-methylpropyl	leu	L
isoleucine	2-butyl	ile	
proline	propyl* - cyclized	pro	Р
phenylalanine	benzyl	phe	F
trytophan	3-indolylmethl		
methionine	methylthioethyl		

Amino Acids - Uncharged Polar R Groups

Amino Acid	Radical	Abbreviations		Abbreviations
		3-Letter	l-Letter	
glycine	Н	gly	G	
serine	hydroxymethyl	ser	S	
threonine	1-hydroxyethyl	thr	T	
cysteine	thiolmethyl	cys		
tyrosine	4-hydroxyphenylmethyl	tyr	Y	
asparagine	aminocarbonylmethyl	asn	N	
glutamine	aminocarbonylethyl	gln Q		

20 Amino Acids - Charged R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	l-Letter
aspartic acid	carboxymethyl	asp	D
glutamic acid	carboxyethyl	glu	E
lysine	4-aminobutyl	lys K arg R	
arginine	3-guanylpropyl		
histidine	4-imidazoylmethyl	his H	

14

Similar minor modifications may also include amino acids deletions or insertions or both. Guidance in determining which amino acid residues may be modified as indicated above without abolishing the desired biological 5 functionality may be determined using computer programs well known in the art, for example, DNASTAR software. addition, the derivative may also result from chemical modifications to the encoded polypeptide, including but not limited to the following, replacement of hydrogen by an alkyl, acyl, or amino group; esterification of a carboxyl 10 group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative. Further a derivative may also result from the substitution of a Lconfiguration amino acid with its corresponding D-15 configuration counterpart.

The term "mimetic", as used herein, refers to a mclecule, the structure of which is developed from knowledge of the structure of a protein/polypeptide or portions thereof (such as BAG-1) and, as such, is able to effect some or all of the actions of BAG-1 protein.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al., Anticancer Drug Des. 8:53-63 (1993)).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a family of BAG-1 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1S beginning at residue 116 of SEQ ID NO:2, BAG-2 (SEQ ID

15

NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO: 8; and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate C.elegans [BAG-1 (SEQ ID NO:12), BAG-2 (SEO ID NO:14)] and the fission yeast S.pombe [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)], full length amino acid sequences specifically the comprising human BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), and BAG-2 (SEQ ID NO:4) C. elegans EAG-1 (SEQ ID NO:12), and BAG-2 (SEQ ID NO:14), and 10 S.pombe BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and partial sequences comprising human BAG-3 (SEQ ID NO: 6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) and functional fragments thereof. In particular, the invention provides the amino acid sequences comprising human BAG-2 (SEQ ID 15 NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) proteins.

Another aspect of the present invention provides the nucleic molecule and nucleotide sequences that encode the family of BAG-1 related proteins from humans [BAG-1 (SEQ ID NO:1), BAG-2 (SEQ ID NO:3), BAG-3 (SEQ ID NO:5) and (SEQ ID NO:19), BAG-4 (SEQ ID NO:7) and (SEQ ID NO:21) and BAG-5 (SEQ ID NO:9) and (SEQ ID NO:23)], the invertebrate C.elegans [BAG-1 (SEQ ID NO:11), BAG-2(SEQ ID NO:13)] and the fission yeast S.pombe [BAG-1A (SEQ ID NO:15), BAG-1B (SEQ ID NO:17)].

BAG-1L (SEQ ID NO:2) is a multifunctional protein that blocks apoptosis, promotes tumor cell metastasis, and contributes to factor-independent and p53-resistant cell growth. BAG-1L (SEQ ID NO:2) interacts with several types of proteins, including Bcl-2, some tyrosine kinase growth

16

factor receptors, steroid hormone receptors, and the p53-induced cell cycle regulator Siah-1A.

BAG-1 is a regulator of Hsc70/Hsp70 family molecular chaperones. A carboxyl-terminal domain in this 5 protein binds tightly to the ATPase domains of Hsc70 and Hsp70 ($K_D = 1$ nM) (Zeiner, M., Gebauer, M., and Gehring, U., EMBO J. 16: 5483-5490, (1997)). BAG-1 modulates the activity of these molecular chaperones, acting as an apparent functional antagonist of the Hsp70/Hsc70-10 associated protein Hip (3-5) (Höhfeld, J. and Jentsch, S., EMBO J. 16: 6209-6216, (1997); Takayama, S., Bimston, D. N., Matsuzawa, S., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R. J., and Reed, J. C., EMBO J. 16: 4887-96, (1997); Zeiner, M., Gebauer, M., and Gehring, U., EMBO J. 16: 5483-5490, (1997)). In general, protein refolding is 15 accomplished by Hsp70/Hsc70 through repeated cycles of target peptide binding and release, coupled to ATP hydrolysis (Ellis, R., Curr Biol. 7: R531-R533, (1997)). BAG-1 appears to promote substrate release, whereas Hip stabilizes Hsp70/Hsc70 complex formation with target 20 peptides (Höhfeld, J., Minami, Y., and Hartl, F.-U., Cell. **83**: 589-598, (1995)). Since each substrate interaction with Hsc70/Hsp70 is unique in terms of the optimal length of time the protein target should remain complexed with 25 Hsc70/Hsp70 for achieving new conformations, the net effect of BAG-1 can be either enhancement or inhibition of the refolding reaction.

The 70kd heat shock family proteins (Hsp70/Hsc70) are essential to a variety of cellular processes and have been implicated in cancer, yet it is unclear how these proteins are regulated *in vivo*. A variety of co-chaperones have been identified which may target Hsp70/Hsc70 to different subcellular compartments or promote their

ANSDOOD -WO

17

interactions with specific protein or protein complexes. BAG-1 appears to represent a novel Hsp70/Hsc70 regulator which differs functionally from all other mammalian cochaperones identified to date, such as members of the DnaJ-, Hip-, Hop-, and cyclophilin-families of proteins.

Another aspect of the present invention provides the amino acid sequence of a binding domain of about 40 to 55 amino acids that bind the a Hsc70/Hsp70 ATPase domain. The BAG domain is situated near the C-terminus, and the ubiquitin-like domains are situated near the N-terminus.

The BAG family of proteins of the present invention contain a common conserved C-terminal domain (the "BAG" domain) that facilitates binding to the ATPase domain of Hsp70/Hsc70. The carboxyl-terminal domain of BAG-1 binds to the ATPase domain of Hsc70/Hsp70 and regulates its chaperone function by acting as a ADP-ATP exchange factor. Other domains of BAG-1 mediate interactions with proteins such as Bcl-2 and retinoic acid receptors (RARs), allowing BAG-1 to target Hsc70/Hsp70 to other proteins, presumably modulating their function by changing their conformations.

Human BAG-1 was previously shown to inhibit Hsc70/Hsp70 dependent refolding of denatured protein substrates in vitro (S. Takayama, et al., EMBO J 16, 4887-96 (1997); M. Zeiner, M. Gebauer, U. Gehring, EMBO J. 16, 5483-5490 (1997); and J. Höhfeld, S. Jentsch, EMBO J. 16, 6209-6216 (1997)). In Example III, Part A the effects of recombinant human BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were compared using in vitro protein refolding assays similar to those employed previously for assessing BAG-1. The study showed that addition of equimolar amounts of each of the recombinant proteins to Hsc70 resulted in significant inhibition of luciferase refolding, with BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) showing somewhat

10

15

greater inhibitor activity than BAG-1 (Figure 4A). In a separate luciferase folding study BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) once again displayed inhibition of luciferase refolding, however in this study varying amounts of BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were added relative to Hsc70 which resulting in concentration-dependent inhibition of Hsc70 chaperone activity, i.e., luciferase folding (Example III Part A). Additional follow on studies using the same experimental protocols as the previous studies, as taught in Example IIA, have shown that BAG-4 (SEQ ID NO:22) also undergoes association with Hsc70/ATPase.

Yet another aspect of the present invention provides a nucleotide sequence having at least about 15 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides that can hybridize or is complementary under relatively stringent conditions to a portion of the nucleic acid sequences shown in Figures 1-9 and Figures 15-17, in particular the BAG domain as shown in in Figure 1B, e.g., nucleotides 552-593 of human BAG-3, or nucleotides 167-221 of human BAG-4.

Yet another aspect of the present invention 25 provides a compound of the formula,

$$R^{N}-R^{1}X^{1}R^{2}X^{2}R^{3}X^{3}R^{4}X^{4}R^{5}X^{5}R^{6}X^{6}R^{7}X^{7}X^{8}R^{9}X^{9}R^{10}X^{10}R^{11}X^{11}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{10}R^{11}X^{10}-R^{C}X^{$$

wherein,

 $\mathbb{R}^{\mathbb{N}}$ is a group of 1 to 552 independently selected amino acids;

R¹ is a group of 3 independently selected amino acids;

 ${\rm X}^1$ is an amino acid with a charged or uncharged R group, such as aspartic acid, glutamic acid, asparagine, or glutamine;

 R^2 is a group of 7 independently selected amino

5 acids;

 χ^2 is an amino acid with a charged R group, such as glutamic acid;

 ${\ensuremath{\mbox{R}}}^3$ is a group of 5 independently selected amino acids;

 x^3 is an amino acid with an apolar R group, such as leucine, methionine, or isoleucine;

 \mathbb{R}^4 is a group of 3 independently selected amino acids;

 χ^4 is an amino acid with charged R group, such as aspartic acid or glutamine acid;

 R^5 is a single independently selected amino acid; X^5 is an amino acid with apolar or uncharged R group, such as leucine, valine, methionine, alanine or threonine;

20 R⁶ is a group of 15 independently selected amino acids;

 $\rm X^6$ is an amino acid with a charged or uncharged R group, such as arginine, lysine, glutamine or aspartic acid;

25 R⁷ is a group of 2 independently selected amino acids;

 χ^7 is an amino acid with a charged R group, such as arginine;

 x^8 is an amino acid with a charged R group, such as arginine or lysine;

 ${\ensuremath{\mbox{R}}}^9$ is a group of 2 independently selected amino acids;

 χ^9 is an amino acid with an apolar R group, such as valine;

 R^{10} is a group of 3 independently selected amino acids;

20

 ${\rm X}^{10}$ is an amino acid with an uncharged R group, such as glutamine;

R¹¹ is a group of 2 independently selected amino acids;

 X^{11} is an amino acid with an apolar R group, such as leucine; and

 $\ensuremath{\,\text{R}^{\text{C}}}$ is a group of 1 to 100 independently selected amino acids.

A nucleotide sequence of at least about 15 nucleotides and, generally, about 25 10 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides can be useful, for example, as a primer for the polymerase chain reaction (PCR) or other similar reaction mediated by 15 a polymerase such as a DNA or RNA polymerase (see PCR Protocols: A guide to methods and applications, ed. Innis et al. (Academic Press, Inc., 1990), which is incorporated herein by reference; see, for example, pages 40-41). In addition, such a nucleotide sequence of the invention can be useful as a probe in a hybridization reaction such as 20 Southern or northern blot analysis or in a binding assay such as a gel shift assay.

A nucleotide sequence of the invention can be particularly useful as an antisense molecule, which can be DNA or RNA and can be targeted to all or a portion of the 5'-untranslated region or of the 5'-translated region of a bag-1 nucleic acid sequence in a cell. For example, an antisense molecule can be directed to at least a portion of the sequence shown as the BAG domain in Figure 1A, e.g., nucleotides 272-319 of human BAG-1L (SEQ ID NO:1), or nucleotides 79-147 of human BAG-5 (SEQ ID NO:9). Since the 5'-region of a nucleic acid contains elements involved in the control of expression of an encoded protein, an antisense molecule directed to the 5'-region of a nucleic

21

acid molecule can affect the levels of protein expressed in a cell.

A nucleotide sequence of the invention also can be useful as a probe to identify a genetic defect due a mutation of a gene encoding a BAG protein in a cell. Such a genetic defect can lead to aberrant expression of a BAG protein in the cell or to expression of an aberrant BAG protein, which does not properly associate with a Bcl-2-related protein or Hsc70/Hsp70 protein in the cell. As a result, a genetic defect in a gene encoding, for example, human BAG-1 can result in a pathology characterized by increased or decreased levels in protein folding.

Further a nucleotide compound or composition as taught in the present invention can be synthesized using routine methods or can be purchased from a commercial 15 In addition, a population of such nucleotide sequences can be prepared by restriction endonuclease or mild DNAse digestion of a nucleic acid molecule that contains nucleotides as shown in the nucleotide sequences shown in Figures 1-9 and Figures 15-17 that encodes the amino acids sequences also shown in Figures 1-9 Methods for preparing and using such Figures 15-17. nucleotide sequences, for example, as hybridization probes to screen a library for homologous nucleic acid molecules are well known in the art (see, for example, Sambrook et 25 al., Molecular Cloning: A laboratory manual (Cold Spring Harbor Laboratory Press 1989); Ausubel et al., Current Protocols in Molecular Biology (Green Publ., NY each of which is incorporated herein by reference).

A particular nucleotide sequence can be designed based, for example, on a comparison of the nucleic acid molecules encoding any one of the BAG family proteins, as shown in Figures 1-9 and Figures 15-17, with another in the family. Such a comparison allows, for example, the

preparation of a nucleotide sequence that will hybridize to a conserved region present in both nucleic acid molecules, thus providing a means to identify homologous nucleic acid molecules present in other cell types or other organisms. In addition, such a comparison allows the preparation of a nucleotide sequence that will hybridize to a unique region of any of the BAG family nucleotide sequences, such as those corresponding to the BAG domain, thus allowing identification of other proteins sharing this motif. this regard, it is recognized that, while the human BAG-3 proteins shown as Figures 3 and 20, and human BAG-5 proteins shown as Figures 5 and 24, are only partial sequences, a variant human BAG-3 or BAG-5 produced, for example, by alternative splicing can exist and can be identified using an appropriately designed nucleotide 15 sequence of the invention as a probe. Such useful probes readily can be identified by inspection of the sequences shown in the disclosed Figures by a comparison of the encoding nucleotide sequences.

If desired, a nucleotide sequence of the invention can incorporate a detectable moiety such as a radiolabel, a fluorochrome, a ferromagnetic substance, a luminescent tag or a detectable binding agent such as biotin. These and other detectable moieties and methods of incorporating such moieties into a nucleotide sequence are well known in the art and are commercially available. A population of labelled nucleotide sequences can be prepared, for example, by nick translation of a nucleic acid molecule of the invention (Sambrook et al., supra, 1989; Ausubel et al., supra, 1989).

One skilled in the art would know that a method involving hybridization of a nucleotide sequence of the invention can require that hybridization be performed under relatively stringent conditions such that nonspecific background hybridization is minimized. Such hybridization

23

conditions can be determined empirically or can be estimated based, for example, on the relative GC content of a sequence and the number of mismatches, if known, between the probe and the target sequence (see, for example, Sambrook et al., supra, 1989).

invention further provides antibodies specific for human BAG family protein. As used herein, the "antibody" includes polyclonal and monoclonal antibodies, as well as polypeptide fragments of antibodies that retain a specific binding activity for human BAG-1 of 10 at least about 1 x 10^5 M⁻¹. One skilled in the art would know that anti-BAG-1 antibody fragments such as Fab, F(ab')and Fv fragments can retain specific binding activity for human BAG-1 (beginning at residue 116 of SEQ ID NO:2) and, thus, are included within the definition of an antibody. 15 In addition, the term "antibody" as used herein includes naturally occurring antibodies as well as non-naturally occurring antibodies and fragments that retain binding activity such as chimeric antibodies or 20 antibodies. Such non-naturally occurring antibodies can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains as described by Huse 25 et al., Science 246:1275-1281 (1989), which is incorporated herein by reference.

One skilled in the art would know that purified BAG family protein, which can be prepared from natural sources or synthesized chemically or produced recombinantly, or portions of a BAG family protein, including a portion of human BAG family protein such as a synthetic peptide as described above, can be used as an immunogen. Such peptides useful for raising an antibody include, for example, peptide portions of the N-terminal 85 amino acids or the BAG domain of any of the human BAG

24

proteins (see Figure 1B). A particularly advantageous use of such a protein is for the immunostaining, wherein the methods provides a process to contrast the immunostaining cf BAG-family proteins in carcinoma cells with adjacent 5 non-neoplastic prostatic epithelial and basal cells which are generally present in the same tissue sections. results would be correlated with a Gleason grade to determine whether any of the BAG-family proteins tend to be expressed at higher or lower levels in histologically advanced tumors. From this process a determination can be made as to degree at which the disease is progressing in a given patient, i.e., a prognosis can be made.

Non-immunogenic fragments or synthetic peptides of BAG proteins can be made immunogenic by coupling the 15 hapten to a carrier molecule such bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH), as described in Example IV, below. In addition, various other carrier molecules and methods for coupling a hapten to a carrier molecule are well known in the art and described, for 20 example, by Harlow and Lane, Antibodies: A laboratory manual (Cold Spring Harbor Laboratory Press, 1988), which is incorporated herein by reference.

EXAMPLES

The following examples are given to enable those 25 skilled in the art to more clearly understand and to practice the present invention. They should not be considered as limiting the scope of the invention, but merely as being illustrative and representative thereof.

25

EXAMPLE I

Isolation and Characterization of BAG-family cDNA Sequences

This example describes methods for isolating and characterizing of BAG-family cDNA sequences from human, nematode and yeast.

A. Cloning of human BAG cDNA sequences

Yeast two-hybrid library screening of a human Jurkat cell cDNA library was performed as described by Takayama et al., <u>EMBO J.,</u> 16:4887-96 (1997); Matsuzawa et 10 al., EMBO J., 17:2736-2747 (1998), which are incorporated herein by reference) using EGY48 strain yeast transformed with pGilda-Hsc70/ATPase (67-377 amino acids) and the lacZ reporter plasmid pSH18-34. Of the resulting $^{\circ}5$ x 10^{6} 15 transformants, 112 Leu colonies were obtained after 1 week incubation at 30° C. Assay of β -galactosidase (β -gal) activity of these colonies resulted in 96 clones. Mating tests were then performed using RFY206 yeast strain transformed with pGilda, pGilda mBAG-1 (1-219), or pGilda 20 Hsc70/ATPase. Of these, 66 displayed specific interactions with Hsc70/ATPase. The pJG4-5 cDNAs were recovered using KC8 E. coli strain which is auxotrophic for tryptophan DNA sequencing revealed 3 partially overlapping human BAG-1, 4 identical and one overlapping cDNAs encoding BAG-2, and 2 partially overlapping BAG-3 clones.

Using the above described yeast two-hybrid screen with the ATPase domain of Hsc70 as "bait", several human cDNAs were cloned which encode portions of BAG-1 or of two other BAG-1-like proteins which are termed BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6). The longest of the cDNAs for BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) contained open reading frames (ORFs) of 207 and 162 amino acids, respectively, followed by stop codons. All BAG-1 (SEQ ID

26

NO:1), BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) cDNAs obtained by two-hybrid library screening with Hsc70/ATPase contained a conserved domain of about 40-50 amino acids which are termed the "BAG" domain and are shown in Figure 10. These results demonstrate that a family of BAG-1-related proteins all contain a conserved ~45 amino acid region near their C-terminus that binds Hsc70/Hsp70.

B. Identification of additional BAG-family proteins

A search of the translated Genbank database using the bBLAST and FASTA search programs also identified human ESTs that provided sequences for further investigation of BAG-family proteins. The putative BAG-4 (SEQ ID NO:8) and BAG-5 (SEQ ID NO:10) proteins contain BAG-domains that share the greatest sequence similarity with the BAG-domain of BAG-3 (SEQ ID NO:6). These were designated BAG-4 (Accession number AA693697, N74588) and BAG-5 (Accession number AA456862, N34101). BAG-4 has 62% identity and ~81% similarity to BAG-3, and BAG-5 has 51% identity and ~75% similarity to BAG-3.

Additional BAG-family orthologues or homologues were also identified using computer-based searches and resulted in BAG-family homologue in the nematode *C. elegans* and the fission yeast *S. pombe*. The *C. elegans* genome encodes two apparent BAG-family proteins, which are most similar in their overall sequences to the human BAG-1 (Afo39713, gi:2773211) (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) (Afo68719, gi:3168927). The *S. pombe* contains two BAG-family proteins that share the greatest overall sequence similarity with human BAG-1 (Alo23S54,gi/3133105 and Alo23634, gi/3150250). The human and *C. elegans* BAG-1 proteins as well as *S. pombe* BAG-1A all have ubiquitin-like domains near their N-termini (see Figure 10A) of unknown function.

The overall predicted amino acid sequences of the C. elegans BAG-1 (SEQ ID NO:12) and S. pombe BAG-1A (SEQ ID NO:16) proteins are ~18% identical (~61% similar) and ~17% identical (~64% similar), respectively, to human BAG-1, implying origin from a common ancestral gene. elegans BAG-1 protein (SEQ ID NO:12), however, contains a 5 to 7 amino acid insert in its BAG-domain as compared to the human, murine, and yeast BAG-1 homologues (see Figure 10B), and is more similar to BAG-2 (SEQ ID NO:4) in regard to its BAG-domain. C. elegans and human BAG-2 also may be 10 derived from a common ancestor as the C-terminal 225 amino acid region which encompasses both the BAG domain and upstream region of both C. elegans and human BAG-2 share ~34% amino acid sequence identity and ~70% similarity. The 15 human BAG-2 protein (SEQ ID NO:4), however, contains a 9 amino acid insert in its BAG-domain compared to it C.elegans counterpart (see Figure 10B). Evolutionary-tree prediction algorithms suggest that human and C. elegans BAG-2 represent a distinct branch of the BAG-family that is more evolutionarily distant from the other BAG-family None of the predicted BAG-family proteins proteins. contain recognizable regions analogous to those found in other Hsc70 regulatory proteins, such as the J-domains and family proteins G/F-domains of DnaJ Tetratricopeptide Repeat (TR) domains of Hip/Hop family 25 proteins.

C. Yeast two-hybrid assay of BAG binding to Hsc70/ATPase

The longest of the cDNAs obtained for the BAG-2 and BAG-3 proteins were expressed with N-terminal transactivation (TA) domains in yeast and tested by yeast two-hybrid assay for interactions with fusion proteins consisting of Hsp70/ATPase or a variety of unrelated proteins (Fas, Siah, Fadd) containing N-terminal LexA DNA-binding domains. TA-BAG-2 and TA-BAG-3 demonstrated

28

positive interactions with LexA-Hsc70/ATPase, resulting in transactivation of a lacZ reporter gene that was under the control of LexA operators (Figure 11A). No interactions with LexA-Fas (cytosolic domain), LexA-Siah, LexA-Fadd, or 5 LexA were detected (see Figure 11A) demonstrating that the BAG-2 and BAG-3 proteins interact specifically with Specific two-hybrid interactions between Hsc70/ATPase. Hsc70/ATPase and either BAG-2 or BAG-3 were also observed when BAG-2 and BAG-3 were expressed as LexA DNA-binding 10 domain fusion proteins and Hsc70/ATPase was fused with a TA domain (see Figure 11A; right panel). These results demonstrate that similarly to BAG-1, BAG-2 and BAG-3 specifically interact with Hsc70/ATPase.

In order to determine whether the BAG proteins are capable of forming heterodimers, coexpression of BAG-2 and BAG-3 in the yeast two-hybrid assay was also performed. Coexpression of BAG-2 and BAG-3 failed to show interaction with BAG-1 or a deletion mutant of BAG-1 (ΔC) which is missing part of its C-terminal domain required for Hsp70/Hsc70 binding suggest that these proteins do not form heterdimers.

D. Isolation and characterization of the complete open reading frame sequences of BAG-2 and BAG-3

In order to deduce the complete ORFs of BAG-2 and BAG-3, a λ -phage cDNA library was screened as follows, using hybridization probes derived from the two-hybrid screening. A human jurkat T-cell λ -ZapII library cDNA library (Stratagene) was screened by hybridization using \$^{32}P\$-labeled purified insert DNA from the longest of the human BAG-2 (clone #11) and human BAG-3 (clone #28) cDNA clones. From about one million clones screened, 38 BAG-2 and 23 BAG-3 clones were identified, cloned, and their cDNA inserts recovered as pSKII plasmids using a helper phage method (Stratagene). DNA sequencing of λ -phage derived

29

human BAG-2 cDNA clones revealed an ORF encoding a predicted 211 amino acid protein, preceded by an in-frame stop codon. The longest human BAG-3 λ -phage cDNA clone contains a continuous ORF of 682 amino acids followed by a stop codon, but without an identifiable start codon (see Figure 10A).

Although BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) all contain a homologous BAG domain near their C-terminus, the N-terminal regions of these proteins are dissimilar. Using a combination of search tools (Prosite Search: PP search, using the Prosite pattern database, BCM Search Launcher, Baylor College of Medicine, and Blocks Search), it was determined that the BAG-2 N-terminal region contains potential kinase phosphorylation sites but otherwise shares no apparent similarity with other proteins or known functional domains.

In contrast, the predicted N-terminal region BAG-3 contains a WW domain as shown in Figure 10A. 20 domains have been identified in a wide variety of signaling proteins, including a Yes kinase adaptor protein (YAP), the regulator Nedd4, Na - channel formin-binding proteins, dystrophin, and the peptidyl prolyl cis-trans-isomerase Pin-1. These roughly 40 amino acid domains mediate protein 25 interactions and bind the preferred peptide ligand sequence xPPxY (Sudol., TIBS, 21: 161-163, 1996, which incorporated herein by reference).

EXAMPLE II

In vitro Association of BAG proteins and Hsc70/ATPase

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) bind Hsc70/ATPase in various in vitro assays.

A. Solution binding assay of BAG-2 and BAG-3 to Hsc70/ATPase

Association of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ 10 ID NO:6) with Hsc70/ATPase was determine by an in vitro protein binding assay where Hsc70/ATPase or BAG-family proteins were expressed in bacteria as Glutathione S-Transferase (GST) fusion proteins. Purified cDNA sequences encoding residues 5 to 211 of human BAG-2 (clone #11) and 15 the C-terminal 135 amino acids of human BAG-3 (clone #28) (see Figure 10A) were subcloned into the EcoRI/Xho I sites of pGEX4T-1 prokaryotic expression plasmid (Pharmacia; Piscataway, NJ). These plasmids as well as pGEX4T-1-BAG-1, pGEX-4T-1-BAG-1 (ΔC), and pGEX-4T-1-XL which have been 20 described previously (Takayama et al., supra (1997); Xie et Biochemistry, 37:6410-6418, (1998), which incorporated herein by reference), were expressed in XL-1 blue strain E. Coli (Stratagene, Inc., La Jolla, Briefly, a single colony was inoculated into 1L of LB media 25 containing 50 μ g/ml ampicillin and grown at 37°C overnight. culture was then diluted by half with LB/ampicillin and cooled to room temperature for 1 hr, before inducing with 0.4mM IPTG for 6 h at 25°C.

Cells were recovered and incubated with 0.5 mg/ml
30 lysozyme in 50 mM Tris (pH 8.0), 150 mM NaCl, 1% Tween-20,
0.1% 2-mercaptoethanol, 5 mM EDTA, 1 mM PMSF and a mixture

of other protease inhibitors obtained from Boehringer Mannheim (1697498) at room temperature for 0.5 h, followed sonication. Cellular debris were pelleted centrifugation at 27,500g for 10 min and the resulting 5 supernatants were incubated with 30 ml of glutathionine-Sepharose (Pharmacia) at 4° C overnight. The resin was then washed with 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, and 0.1% 2-mercaptoethanol until the OD 280nm reached <0.01. For removal of GST, the resin with immobilized GST-10 fusion protein was incubated with 10U (Boehringer, Inc.) at 4° C in 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, 0.1% 2-Mercaptoethanol, and 2.5 mM CaCl2 overnight. Released proteins were then purified on Mono Q (HR10/10, Pharmacia) by FPLC using a linear gradient 15 of 0.5M NaCl at pH 8.0 and dialyzed into chaperone assay buffer.

The ability of BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ NO:6) bind Hsc70/ATPase in solution was then examined. GST control or GST-BAG proteins were immobilized on glutathione-Sepharose and tested for binding to 35Slabeled in vitro translated (IVT) proteins. Immunoprecipitation and in vitro GST-protein binding assays were performed as described by Takayama et al., supra (1997), using pCI-Neo flag or pcDNA3-HA into which human Bag-2 (clone #11) or human BAG-3 (clone #28) had been 25 subcloned for in vitro translation of 35S-L-methionine labeled proteins or expression in 293T cells. As shown in Figure 11B, 35S-Hsc70/ATPase bound in vitro to GST-BAG-1, GST-BAG-2, and GST-BAG-3 but not to GST-BAG-1(Δ C) or 30 several other control proteins. BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) also exhibited little or no binding to themselves or to each other, demonstrating that these proteins do not strongly homo- or hetero-dimerize or 35 oligomerize. It should be noted, however, that BAG-2 (SEQ

ID NO:4) displayed weak interactions with itself in binding assays and produced a positive result in yeast two-hybrid experiments, demonstrating that it can have the ability to self-associate.

5 B. Binding of BAG proteins to Hsc70 in vivo

The ability of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins to interact in cells with Hsc70 was tested by expressing these proteins with N-terminal Flag epitope tags in 293T human epithelial cells using coimmunoprecipitation assays as described 10 (Takayama et al., supra (1997)). cDNAs encoding the λ phage cloned regions of BAG-2 and BAG-3 were subcloned inframe into pcDNA3-Flag. Anti-Flag immune complexes prepared from 293T cells after transfection with plasmids encoding Flag-BAG-1, Flag-BAG-2, or Flag-BAG-3 15 analyzed by SDS-PAGE/immunoblot assay. As shown in Figure 10C, antiserum specific to Hsc70 detected the presence of BAG proteins associated with Hsc70, whereas control immunecomplexes prepared with IqG1 as well as anti-Flaq immune 20 complexes prepared from cells transfected with Flag-tagged control proteins, Daxx and Apaf-1, did not contain Hsc70 associated protein. These results further demonstrate that BAG-family proteins specifically bind to Hsc70.

C. BIAcore assay of BAG protein binding to the ATPase domain of Hsc70

BAG-1 (beginning at residue 116 of SEQ ID NO:2) is known to bind tightly to the ATPase domain of Hsc70 (Stuart et al., <u>J. Biol. Chem.</u>, In Press (1998)). BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins were therefore, examined for their ability to bind to Hsc70/ATPase. The affinity and binding kinetics of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) to Hsc70/ATPase was also compared to that of BAG-1 (beginning at residue 116 of

33

SEQ ID NO:2) for Hsc70/ATPase, using a surface plasmon resonance technique (BIAcore) which has been described previously (Stuart et al., supra, (1998) which is incorporated herein by reference).

5 BAG-family proteins were produced in bacteria and purified to near homogeneity as shown in Figure 12A and described above in Example I. The purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), -2 (SEO ID NO:4), and -3 (SEQ ID NO:6) proteins were then immobilized 10 on biosensor chips and tested for their interactions with Hsc70 in the soluble phase. Kinetic measurements were performed using a BIAcore-II instrument with CM5 sensor chip and Amine Coupling Kit (Pharmacia Biosensor AB, Sweden). Briefly, for immobilization of proteins, the sensor chip was equilibrated with HK buffer (10 mM Hepes 15 (pH 7.4), 150 mM KCL) at 5μ l/min, then activated by injecting 17 μ l of 0.2M N-ethyl-N'-(3-diethylaminopropyl)carbodiimide and 0.05M N-hydroxysuccinimide followed by 35 μ l of the protein of interest, in 10 mM 20 acetate, pH 3.5-4.5. Excess NHS-ester on the surface was deactivated with 17 μ l 1M ethanolamine-HCL (pH8.5). After immobilization, 5μ l of regeneration buffer (50 mM phosphate (pH 6.8) and 4M GuHCl) was injected. For binding assays, Hsp70 (Sigma, H8778) was dissolved in HK buffer, injected at 10 μ l/min across the prepared surface at 25 various concentrations. The surface was regenerated after each injection with 5 μ l of regeneration buffer. The rate constants κ_{ass} and κ_{diss} were generated with BIAevaluation softward 3.01 (Pharmacia Biosensor AB). Addition of Hsc70 to chips containing BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) resulted in concentration-dependent binding, as reflected by an increase in the Response Units (RU) measured at the chip surface (shown in Figure 3B). In contrast, Hsc70 failed to display interactions in BIAcore assays with a 35 variety of control proteins as well as a mutant of BAG-1

34

lacking a C-terminal portion of the BAG domain which is required for Hsc70-binding (Figure 3B). Furthermore, flowing of various control proteins such as GST, BSA and Bcl-XL over the BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) chips resulted in negligible interaction. These results further demonstrate the specificity with which BAG-family proteins interact with and bind to Hsc70.

The rates of Hsc70 binding to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 10 (SEO ID NO:6) proteins were similar, following pseudo first-order kinetics with estimated association rate constants (κ_a) of 2.1, 2.1 and 2.4 x 10^5 M⁻¹ sec⁻¹, respectively. After allowing binding of Hsc70 to immobilized BAG-1 (beginning at residue 116 of SEQ ID 15 NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to reach plateau levels, the chaperone was removed from the flow solution and the dissociation rate was monitored. BAG-1 (beginning at residue 116 at SEQ ID NO:2) and BAG-2 (SEQ ID NO:4) exhibited similar dissociation rates, with relatively 20 slow loss of Hsc70 from the chip surface, resulting in estimated dissociation rate constants (κ_d) of 3.0 and 5.0 x 10⁻⁴ sec⁻¹, respectively (see Figure 3B). In contrast, Hsc70 dissociated more rapidly from biosensor chips containing BAG-3 (see Figure 3B), yielding an estimated κ_d of 1.7 x 10⁻³ 25 sec^{-1} . From the kinetic data, the apparent affinities (κ_{D} = κ_d/κ_a) were calculated for binding of Hsc70 to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) and were estimated to equal about $K_D=1.4 \text{nM}$, $K_D=2.4 \text{nM}$, and $K_D=7.4 \text{nM}$, respectively. These results demonstrate that the interactions of BAG-family proteins with Hsc70 occur with apparent affinities sufficient for physiological relevance.

35

EXAMPLE III

BAG-family proteins inhibit Hsp70/Hsc70-dependent protein folding

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins inhibit Hsp70/Hsc70-dependent refolding of denatured proteins similarly to a BAG-1 (beginning at residue 116 of SEQ ID NO:2) protein.

The effects of BAG-2 (SEQ ID NO:4) and BAG-3 (SEO ID NO:6) protein on Hsp70/Hsc70-dependent protein refolding 10 was determined using in vitro protein refolding assays similar to those described previously by Takayama et al., supra, 1998: Terada et al., <u>J Cell Biol.</u>, 139:1089-1095, 1997, which are incorporated herein by reference. Briefly, luciferase (20 μ M) was denatured in 25 mM Hepes-KOH, pH 7.2, 15 50 mM potassium acetate, 5 mM DTT, hydrochloride at ~25°C for 1 h. Denatured luciferase was diluted 1:40 into 25 mM Hepes-KOH, pH 7.2, 50 mM potassium acetate, 5 mM DTT. Hsc70 (1.8 μ M), DnaJ (StressGen, Inc.) $(0.9\mu\text{M})$, and various purified recombinant proteins as indicated were added to refolding buffer (30 mM Hepes-KOH, 20 pH 7.6, 120 mM potassium acetate, 3mM magnesium acetate, 2 mM DTT, 2.5 mM ATP) with 0.2 volume of diluted denatured luciferase to a final concentration of 0.1 μ M. Luciferase activity was measured after 1.5 hr incubation at 35°C.

The combination of Hsc70 and DnaJ resulted in ATP-dependent refolding of chemically denatured firefly luciferase, with function of over half the denatured enzyme restored in a 90 minute reaction, as monitored by a chemiluminescence assay. In contrast, neither Hsc70 nor DnaJ alone were able to induce substantial refolding of denatured luciferase. Furthermore, little spontaneous

36

restoration of luciferase activity was observed with control proteins, BSA, GST or Bcl-XL (see Figure 4A).

Addition of recombinant purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to the above assays in amounts equimolar to Hsc70 (1.8 μ M) resulted in striking inhibition of luciferase refolding. BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) displayed somewhat greater inhibitory activity than BAG-1 (beginning at residue 116 of SEQ ID NO:2) as shown in Figure 4A. In contrast, the BAG-1 (Δ C) protein, which fails to bind Hsc70 as well as several other control proteins, had no effect on luciferase refolding.

In an additional refolding assay, described previously by Minami et al., J Biol. Chem. 271:19617-24, 15 1996), purified Hsc70 and human DnaJ homolog Hdj-1 (Hsp 40) with additional cofactors provided reticulocyte lysates (5% v:v) to produce a system capable of refolding denatured luciferase. Briefly, additional cofactors included, recombinant Luciferase (Promega: 20 QuantiLum TM), that had been heat denatured at 42°C for 10 min, 1.8 μM Hsc70 (Sigma; purified from bovine brain), 0.9 μM Hsp40, and various recombinant purified proteins. Luciferase activity was measured (Promega luciferase assay kit) using a luminometer (EG&G Berthold, MicroLumat luminometer, Model #LB96P). All results were normalized relative to non-denatured luciferase that had been subjected to the same conditions. Control reactions lacking ATP, Hsc70, or Hsp40 resulted in negligible luciferase refolding. 30

Various amounts of purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6), relative to amounts of Hsc70 were used in the above-described protein refolding assay. Addition of BAG-family proteins resulted in a concentration-dependent

37

inhibition of Hsc70 chaperone activity. Furthermore, the BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) inhibition of Hsc70 chaperone activity was demonstrated to be as potent as that observed for BAG-1 (beginning at residue 116 of SEQ ID NO:2). In contrast, the BAG-1 (ΔC) mutant as well as other control proteins did not suppress Hsc70-mediated refolding of denatured luciferase. These results indicate that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) can inhibit Hsc70/Hsp70 dependent protein refolding activity to the same extent as BAG-1 (beginning at residue 116 of SEQ ID NO:2).

B. BAG competes with Hip for binding to Hsc70.

It is known that BAG-1 competes with Hip for binding to Hsc70, with these proteins exerting opposite effects on Hsc70-mediated protein refolding (Hohfeld, J., and Jentsch, S., Embo J., 16:6209-6216, 1997, which is incorporated herein by reference). In order to determine whether BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) also compete with Hip for binding to Hsc70, refolding assays were performed as described above in the presence of Hip protein.

Hip was purified as His₆-protein. The fusion protein was induced from pET28-Hip (V. Prapapanich et al., Mol Cell Biol., 18:944-952, 1998, which is incorporated herein by reference) with 0.1 mM IPTG at 25°C for 6h in BL21 cells. Cells from 1L of culture were resuspended into 50 ml of 50 mM Phosphate buffer (pH 6.8), 150 mM NaCl, and 1% (v/v) Tween-20 and then incubated with 0.5 mg/ml lysozyme at 25°C for 0.5h, followed by sonication. After centrifugation at 27,500g for 10 min, the resulting supernatant was mixed with 15 ml nickel resin (Qiagen, Inc.) at 4°C for 3 h with 25 mM imidazol. The resin was then washed with 50 mM phosphate buffer (pH 6.8), 25 mM imidazol, 150 mM NaCl and 0.1% Tween-20 until the OD280nm

38

reached a value of <0.01. His,-Hip protein was eluted with 250 mM imidazol in washing buffer (Qiagene, Inc.) and purified on Mono Q (HR10/10 Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0, followed by dialysis in chaperone assay buffer.

In the refolding assay reactions, addition of purified Hip at equimolar concentrations relative to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) (1.8 µM) completely negated the inhibitory effects of the BAG-family proteins on refolding of denatured luciferase (see Figure 4C). These results demonstrate that the suppression of Hsc70 chaperone activity by BAG-family proteins is reversible, and that Hip antagonizes the effects of not only BAG-1 (beginning at residue 116 of SEQ ID NO:2), but also of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6).

In summary, these results demonstrate that BAG-family proteins all contain a conserved BAG domain near their C-terminus that binds Hsc70/Hsp70, and that human BAG-family proteins can bind with high affinity to the ATPase domain of Hsc70 and inhibit its chaperone activity through a Hip-repressable mechanism.

EXAMPLE IV

EXPANDED NUCLEIC ACID AND AMINO ACID SEQUENCES FOR HUMAN BAG-3, BAG-4 AND BAG-5

Following the procedures disclosed herein, the nucleic acid and amino acids sequences to human BAG-3, BAG-4 and BAG-5 were further expanded. The expanded sequences for BAG-3, BAG-4 and BAG-5 are shown in Figures 15, 16 and 17, respectively, with their respective sequence identification numbers, "SEQ ID NO"s.

We claim:

1. A compound of the formula,

$R^{N} - R^{1}X^{1}R^{2}X^{2}R^{3}X^{3}R^{4}X^{4}R^{5}X^{5}R^{6}X^{6}R^{7}X^{7}X^{8}R^{9}X^{9}R^{10}X^{10}R^{11}X^{11} - R^{C}$

wherein	·
where the	R^N is a group of about 1 to 552 independently
	selected amino acids;
	R ¹ is a group of 3 independently selected amino
	acids;
	\mathbf{X}^{1} is an amino acid with a charged or uncharged
	R group;
	R^2 is a group of 7 independently selected amino
	acids;
	$ exttt{X}^2$ is an amino acid with a charged R group;
	R^3 is a group of 5 independently selected amino
	acids;
	$ exttt{X}^3$ is an amino acid with an apolar R group;
	R ⁴ is a group of 3 independently selected amino
	acids;
	$ exttt{X}^4$ is an amino acid with charged R group;
	R ⁵ is a single independently selected amino acid;
	${\tt X}^{\tt 5}$ is an amino acid with apolar or uncharged R
	group;
	R ⁶ is a group of 15 independently selected amino
	acids;
	${ t X}^6$ is an amino acid with a charged or uncharged
	R group;
	R ⁷ is a group of 2 independently selected amino
	acids;
	$ exttt{X}^7$ is an amino acid with a charged R group;
	$ exttt{X}^8$ is an amino acid with a charged R group;
	R ⁹ is a group of 2 independently selected amino
	acids;
	$ exttt{X}^9$ is an amino acid with an apolar R group;
	wherein,

PCT/US99/21053

5

 \mathbb{R}^{10} is a group of 3 independently selected amino acids;

 \mathbf{X}^{10} is an amino acid with an uncharged R group; \mathbf{R}^{11} is a group of 2 independently selected amino acids;

 \mathbf{X}^{11} is an amino acid with an apolar R group; and \mathbf{R}^{C} is a group of about 1 to 100 independently selected amino acids.

- 2. A substantially purified nucleic acid molecule having a nucleotide sequence corresponding to or complementary to at least 20 nucleotides from a nucleotide sequence selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:21) and (SEQ ID NO:23).
- 3. The nucleic acid of claim 2 having a nucleotide sequence corresponding to or complementary to a nucleotide sequence that encodes a functionally active BAG family protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:24).
- 4. The nucleic acid of claim 3 selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).
 - 5. The nucleic acid of claim 3 complementary to a nucleotide sequence that encodes a functionally active BAG protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).
 - 6. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:3).

41

- 7. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:5).
- 8. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:7).
- 9. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:9).
- 10. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:19).
- 10 11. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:21).
 - 12. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:23).
- 13. A substantially purified BAG family protein encoded by the nucleic acid molecule of claim 1.
 - 14. A substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24) or a fragment, a derivative or a mimetic thereof.
 - 15. A substantially purified protein corresponding to the amino acid sequence of 157 to 204 of (SEQ ID NO:2).
- 25 16. A substantially purified protein corresponding to the amino acid sequence of 272 to 319 of (SEQ ID NO:2).

PCT/US99/21053

- 17. A substantially purified protein corresponding to the amino acid sequence of 164 to 211 of (SEQ ID NO:4).
- 18. A substantially purified protein 5 corresponding to the amino acid sequence of 418 to 510 of (SEQ ID NO:20).
 - 19. A substantially purified protein corresponding to the amino acid sequence of 378 to 457 of (SEO ID NO:22).
- 10 20. A substantially purified protein corresponding to the amino acid sequence of 6 to 97 of (SEQ ID NO:24).
- 21. A substantially purified protein corresponding to the amino acid sequence of 180 to 257 of 15 (SEQ ID NO:24).
 - 22. A substantially purified protein corresponding to the amino acid sequence of 272 to 349 of (SEQ ID NO:24).
- 23. A substantially purified protein 20 corresponding to the amino acid sequence of 362 to 444 of (SEQ ID NO:24).
 - 24. A pharmaceutical composition comprising a nucleic acid molecule of claim 1 useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.
 - 25. A method of modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function by administering a nucleic acid molecule of claim 1.

- 26. A pharmaceutical composition comprising a substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:22) and (SEQ ID NO:24), or a fragment, a derivative or a mimetic thereof, useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.
- 10 27. A method of modulating tumor cell proliferation by administering a pharmaceutical composition of claim 26.
- 28. A method of modulating cell migration and metastasis by administering a pharmaceutical composition of claim 26.
 - 29. A method of modulating steroid hormone receptor function by administering a pharmaceutical composition of claim 26.
- 30. A substantially purified antibody that 20 specifically binds to a BAG family protein of claim 14.
 - 31. The antibody of claim 30, wherein said antibody is a monoclonal antibody.

44

32. A method for detecting the presence of a BAG family protein in a sample, comprising the steps of:

a. obtaining the sample;

5

15

20

PNSDOCID ZWO

- b. adding to said an antibody of claim 11 under suitable conditions for the binding of said antibody with the BAG family protein; and
- c. detecting said bound BAG family protein.
- 33. A method for detecting the presence of a first nucleic acid molecule that encodes a BAG family protein in a sample, comprising the steps of:
 - a. obtaining the sample;
 - b. adding to said sample a second nucleic acid molecule capable of hybridizing with said first nucleic acid molecule under suitable conditions for the binding of said second nucleic acid molecule with said first nucleic acid molecule; and
 - c. detecting said hybridized first and second nucleic acid molecules.
- 34. A method of determining the risk of metastatic spread of cancer or prognosis of cancer patients 25 by determining the level of expression of a BAG-family protein.

ACGCCGCGCT CAGCTTCCAT CGCTGGGCGG TCAACAAGTG CGGGCCTGGC TCAGCGCGGG GGGGCGCGGA GACCGCGAGG CGACCGGAGG L A Q R G G A R R R R G D R E BAG-1L GGGTGGGTT CCCGGTGGG CGCCTTCGG CCAGGCCGGG AGCCCCA GTCGGAGGCC CCGGCCCAGG GTGGTCCGC TCCTCTGG	90 180
A L R TGCCAGCGGG A S G	270
BAG-1M accegece ecteraces grocorgere ttorceger gerospecti greeters grochages cetgrofich rgrogogre t r r r s e e r t w s e e e r t l s e e r t w s e e r r	360
CHGRGTGRGG RGGCGRCCR GGGCGRGGRG RTGRRTCGGR GCCRGGRGGT GRCCCGGGRC GRGGRGTCGR CCCGGRGCGR GGRGGTGRCC ${\sf Q}$ ${\sf S}$ ${\sf E}$ ${\sf E}$ ${\sf R}$ ${\sf N}$ ${\sf N}$ ${\sf R}$ ${\sf S}$ ${\sf Q}$ ${\sf E}$ ${\sf U}$ ${\sf T}$ ${\sf R}$ ${\sf D}$ ${\sf E}$ ${\sf E}$ ${\sf S}$ ${\sf T}$ ${\sf R}$ ${\sf S}$ ${\sf E}$ ${\sf E}$ ${\sf U}$ ${\sf T}$	450
RGGGRGGRAPA TGGCGGCRGC TGGGCTCRCC GTGACTGCRA TGRGARGCRC GRCCTTCRTG TTRCCTCCCR GCRGGGCRGC R E E M A A A G L T V T V T H S N E K H D L H V T S Q Q G S	540
нотбяясся оттероват воттерве востем востем востем востеся востительной востительной в в в в в ${\sf R}$ в в ${\sf R}$	1/3 089
TCTCTGRAGG АЯАТGGARAC ACCGTTGTCA GCACTTGGAR TACRAGATGG TTGCCGGGTC RTGTTARTTG GGARARAGAR CAGTCCACAG S L K E M E T P L S A L G I Q D G C R V M L I G K K M S P Q	9 022
САЯСЯССТТС АЯСТАЯЯСЯЯ СТТСЯВОСЯСТСТСТССОСЯСЯ САТИССТССАЯС АСТТСЯЯТАЯ ВСЯССТТЯСТ ${\sf E}$ ${\sf E}$ ${\sf L}$ ${\sf R}$ ${\sf M}$ ${\sf E}$ ${\sf L}$ ${\sf I}$ ${\sf R}$ ${\sf D}$ ${\sf Q}$ ${\sf L}$ ${\sf E}$ ${\sf E}$ ${\sf L}$ ${\sf N}$ ${\sf E}$ ${\sf L}$ ${\sf I}$	8 10
GGARTCCAGC AGGGTTTCT GCCCARGGAT TTGCARGCTG ARGCTCTCTG CARACTTGAT AGGAGAGTAA ARGCCACART AGAGCAGTTT G I Q Q G F L P K D L Q A E A L C K L D R R V K A T I E Q F	006
АТОЯНСЯТСТ ТОСНОСЯВИТ ТОЯССКОСТВО СТОССОСНО ВНЯНИТІТСЯЯ НОЯССТВОЯ НОВСТВОЕТ В НИТИЗАТИТЕЛЯ В ${\sf R}$.	066
СЯGGCATTCC TRGCCGRGTG TGRCRGRGGRACA TCTGCCRGGA GACTGAGCGG CTGCAGTCTA СЯЯАСТТТGC ССТGGCCGAG Q A F L A E C D T V E Q N I C Q E T E R L Q S T N F A L A E	1080
сновнянянно стотостьсе стояньные всесенсся стетьссете тетьвните натапнесть натастень	1170
GGCTGCTGGG GGCAACTGGC CATTTGCCAA ITTTCCTACT CTCACACTGG TTCTCAATGA ААААТАGTGT CTTTGTGATT 1GAGTAAAGC	1260
TCCTRITCIG ITTITCACAA AAAAAAAAA A	1291

06	180	270	360	450	045 05 5/3	9 089	720	8 10
	CCCTGCAGCC GAAGATCAAC K I N	GCTGGAGCTC L E L	AAATAGCCAG N S Q	TGAAGTGTCA E U S	тст66өт6өт L D D	GTTTCARTCC F Q S	TGACAAGGCC D K A	нянсстянся
	CGCGTTG GTGACGGCGA CCCTGCAGG SCTTAGA TGGCTCAGGC GAAGATCAF N A Q A K I N	GCCTGGACCA L D Q	ACAGTATCCA S I Q	CTCTCRCCGT L T U	TCAATAAGTT N K F	TTGATCAGAA D Q K	TTGARARCTC E N S	яттястсттс
	TTGCCCCCGC GGCGTCAGAG GGAGGGCGGG CGCCGCGTTG GTGACGGCGA CCCTGCAGCC GGCCGGTGAC CTCTTGGCTA CCCCGCGTCG GAGGCTTAGA TGGCTCAGGC GAAGATCAAC	TCCTCCTCCA TGGCTGACCG CTCCAGCCGC CTGCTGGAGGA GCCTGGACCA GCTGGAGCTC S S S M A D R S S R L L E S L D Q L E L	AGGGTTGAAG CTTTGAGAGA AGCAGCAACT GCTGTTGAGC AAGAGAAAGA AATCCTTCTG GAAATGATCC ACAGTATCCA AAATAGCCAG R V E A L R E A A T A V E Q E K E I L L E M I H S I Q N S Q	АТСССАНСАЯ И С В Т	6AT6A66T6G D E U U	TTGGGARATG CCARGAGTCA TTTAATGTCG CTCTACAGTG CATGTTCATC TGAGGTGCCA CATGGGCCAG TTGATCAGAA GTTTCAATCC L 6 N A K S H L M S L Y S A C S S E U P H G P U D Q K F Q S	ATAGTARTTG GCTGTGCTCT TGARGATCAG AAGAAAATTA AGAGAAATT AGAGACTCTG CTTAGAAATA TTGAAAACTC TGACAAGGCC	ATCARGCTAT TAGAGCATTC TARAGGAGCT GGTTCCAAAA CTCTGCAACA AAATGCTGAA AGCAGATTCA ATTAGTCTTC AAACCTAAGA I K L L E H S K G A G S K T L Q Q N R E S R F N
	6646666666 CCCCGC6TCG	CTCCAGCCGC S S R	ARTCCTTCTG	ARACCGTTTG N R L	ARGGATTATT R I I	TGAGGTGCCA E U P	RGAGACTCTG E T L	I AAATGCTGAA N A E
	обсотсново стсттоостя	TGGCTGACCG A D R	янсясяннся Е К Е	HTCTGACTGC	HGCRTGCCAC H A T	CATGITCAIC	I AGAGAAGATT	і стстбсяяся · L Q Q
		TCCTCCTCCA S S S M	GCTGTTGAGC A U E Q	GARGARTTAR E E L N	GARTCCCTAR ESLK	CTCTACAGIG	HAGARARITA K K I K	GOTTCCAAAA
	TCCTCCC666 6CCGCC66A6	CTTCTGCCGC F C R	АGCAGCAACT A A T	CGGAGAAAGA G E R	ссносносня О О О	TTTARTGTCG	тояноятсяо Е D Q	TARAGGAGCT K G A
	GCAGCCGCGG TGTCGCGAAG TCCTCCCGGG CAAGGAGCGC TCCACTCGCT GCCGCCGGAG	GCTAAAGCCA ACGAGGGGG CTTCTGCCGC A K A N E G R F C R	CTTTGAGAGA L R E	AGATCAGTGA ISD	TTAGAAACCC R N P	ссенбенбтся К S Н	GCTGTGCTCT C A L	TAGAGCATTC E H S
	осноссосов сановносос	GCTAAAGCCA A K A N	AGGGTTGAAG R U E A	GACATGAGGC AGATCAGTGA CGGAGAAAAGA GAAGAATTAA ATCTGACTGC AAACCGTTTG ATGGGAAGAA CTCTCACCGT TGAAGTGTCA Дом в q is d g e s e e e e n e t a n b e n g b t e t u e u s	S GTAGARACAR ITAGARACC CCAGCAGCAR GARICCTAR ACCAIGCCAC ARGGAITAIT GAIGAGIGG ICAATARGII ICIGGAIGAI U E T I R N P Q Q Q E S L K H R T R I I D E U U N K F L D D	TTGGGAAATG L G N A	ATAGTAATTG	ATCAAGCTAT

3/39

066 1080 1170 1179 6САТТТЯСЯС ЯЯТЯСЯСЯВ БТБТЯЯЯВЯТ СЯТЯЯЯВТЯС ТЯТТТЯЯТТ СЯТЯЯСТЯСТ ТСТТТБТТЯС БТЯТЯЯССЯС ТТЯБТТСЯСЯ ттсябятбя бояваяятятт ссятсявстя теттсявтт тотовянтяяс явяястявся втятттяят тятстятстя AGATTGARIT CITGICITGI ACTAGGATCI AGCATATITC ACTATICIGI GGATGAATAC ATAGITIGIG GGGAAAACAA AGGGGCARAR AGCATGACTG CITITICCTG ICTGGCATGG ARTCACGCAG ICACCTTGGG CATTIAGITI ACTAGARATI CTGATAGTTG GAGATTTTTT **ACGTTCAGCT** CTTTACTGG

FIG. 2B

4/39

GCGGAGCTCC GCATCCAACC A E L R I Q P	CCGGGCCGCG GCCRACTTCT CTGGACTC	GA CCAGAAGTTT CTAGCCGGCC D Q K F L A G Q	AGTTGCTACC TCCCTTTATC	90
TCCTCCTTCC CCTCTGGCAG S S F P S G S	CGAGGAGGCT ATTTCCAGAC ACTTCCAG	CC CTCTCTGGCC ACGTCACCCC P S L A T S P P	CGCCTTTAAT TCATAAAGGT PLIHKG	180
GCCCGGCGCC GGCTTCCCGG A R R R L P G	ACACGTCGGC GGCGGAGAGG GGCCCACG	GC GGCGGCCCGG CCAGAGACTC A A A R P E T R	GGCGCCCGGA GCCAGCGCCC :	270
CGCACCCGCG CCCCAGCGGG R T R A P A G	CAGACCCCAA CCCAGCATGA GCGCCGCC	AC CCACTCGCCC ATGATGCAGG T H S P M M Q U	TGGCGTCCGG CAACGGTGAC :	360
CGCGACCCTT TGCCCCCGG R D P L P P G	ATGGGAGATC AAGATCGACC CGCAGACG	GG CTGGCCCTTC TTCGTGGACC G W P F F V D H	ACAACAGCCG CACCACTACG N S R T T T	450
TGGAACGACC CGCGCGTGCC W N D P R U P	CTCTGAGGGC CCCAAGGAGA CTCCATCO	TC TGCCAATGGC CCTTCCCGGG S A N G P S R E	AGGGCTCTAG GCTGCCGCCT	540
GCTAGGGAAG GCCACCCTGT A R E G H P U	GTACCCCCAG CTCCGACCAG GCTACAT	CC CATTCCTGTG CTCCATGAAG P I P V L H E G	GCGCTGAGAA CCGGCAGGTG A E N R Q U	630
CACCCTTTCC ATGTCTATCC	CCAGCCTGGG ATGCAGCGAT TCCGAAC			720
CGGGGCATGC CAGAAACCAC R G M P É T T	TCAGCCAGAT AAACAGTGTG GACAGGTG	GC AGCGGCGGCG GCAGCCCAGC A A A A A A Q P	CCCCAGCCTC CCACGGACCT P A S H G P	8 10
GAGCGGTCCC AGTCTCCAGC E R S Q S P A	TGCCTCTGAC TGCTCATCCT CATCCTC			900
CACCAGCTCC CGCGGGGGTA H Q L P R G Y	CATCTCCATT CCGGTGATAC ACGAGCAC	AA CGTTACCCGG CCAGCAGCCC N U T R P A A Q	AGCCCTCCTT CCACAAAGCC PSFHKA	3 90
CAGAAGACGC ACTACCCAGC Q k T H Y P A	GCAGAGGGGT GAGTACCAGA CCCACCAC Q R G E Y Q T H Q			080
CCCCTGCGGG CGGCATCCCC P L R A A S P	GTTCAGGTCA TCTGTCCAGG GTGCATCG F R S S U Q G A S	AG CCGGGAGGGC TCACCAGCCA S R E G S P A R	GGAGCAGCAC GCCACTCCAC 1 S S T P L H	170
TCCCCCTCGC CCATCCGTGT S P S P I R U	GCACACCGTG GTCGACAGGC CTCAGCAC	CC CATGACCCAT CGAGAAACTG P M T H R E T A	CACCTGTTTC CCAGCCTGAA 1 P U S Q P E	260
AACAAACCAG AAAGTAAGCC N K P E S K P	AGGCCCAGTT GGACCAGAAC TCCCTCC		TCCGCARAGA GGTGGATTCT 1 R K E U D S	350
AAACCTGTTT CCCAGAAGCC K P V S Q K P	CCCACCTCCC TCTGAGAAGG TAGAGGTC PPPSEKUEU	AA AGTTCCCCCT GCTCCAGTTC K U P P A P U P	C P P P S P	440

GGCCCTTCTG CTGTCCCCTC	TTCCCCCAAG AGTGTGGCTA CAGAAGAGA S P K S V A T E E A	G GGCAGCCCCC AGCACTGCCC A A P S T A P	CTGCAGAAGC TACACCTCCA A E A T P P	1530
AAACCAGGAG AAGCCGAGGC K P G E A E A	TCCCCCAAAA CATCCAGGAG TGCTGAAAG PPK HPG U L K V	T GGAAGCCATC CTGGAGAAGG E A I L E K V	TGCAGGGGCT GGAGCAGGCT Q G L E Q A	1620
GTAGACAACT TTGAAGGCAA U D N F E G K	GAAGACTGAC AAAAAGTACC TGATGATCG	A AGAGTATTTG ACCARAGAGC EYLTKEL	TGCTGGCCCT GGATTCAGTG	1710
GACCCCGAGG GACGAGCCGA	TGTGCGTCAG GCCAGGAGAG ACGGTGTCA	G GAAGGTTCAG ACCATCTTGG K V Q T I L E		1800
ATTGATGTCC CAGGTCAAGT	CCAGGTCTAT GAACTCCAGC CCAGCAACC	T TGARGCAGAT CAGCCACTGC E A D Q P L Q	AGGCARTCAT GGAGATGGGT A I M E M G	1890
GCCGTGGCAG CAGACAAGGG A U A A D K G	CAAGAAAAAT GCTGGAAATG CAGAAGATC K K N A G N A E D F	C CCACACAGAA ACCCAGCAGC H T E T Q Q P		1980
ACTTCAAACC CCAGCAGCAT T S N P S S M	GACAGACACC CCTGGTAACC CAGCAGCAC	C GTAGCCTCTG CCCTGTAAAA	GTCAGACTCG GAACCGATGT	2070
GTGCTTTAGG GATTTTAGTT	GCATGCATTT CAGAGACTTT AGGTCAGTT	G GTTTTGATTA GCTGCTTGGT	ATGCAGTACT TGGGTGAGGC	2 150
AAACACTATA AAGGGCTAAA	AGGGAAAATG ATGCTTTTCT TCAATATTC	T TACTOTTGTA CAATTAANGA	AGTTGCTTGT TGTTTGAGAA	2250
GTTTARCCCC GTTGCTTGTT	CTGCAGCCCT GTCHACTTGG GCACCCCCA	C CACCTGTTAG CTGTGGTTGT	GCACTGTCTT TTGTAGCTCT	2340
	GAGTCAATTA CCCATCACAT AAATATGAA		ATTTTAATGA GATGATTTTC	2430
TTCATCTCAT AATTAAAATA	CCTGACTTTA GAGAGAGTAA AATGTGCCA	S GAGCCATAGG AATATCTGTA	TGTTGGATGA CTTTAATGCT	2520
ACATTTTH	FIG.3			2528

_	_	_	_			_	5/3		_	-	-
8	180	270	360	450	540	630	720	8 10	306	366	10 10
S GRATTGCARG GCCRGRGIII GRATICITRI RCRARIGGRG CGIRIGGICC RACRIRCCCC CCRGGCCCTG 90	CAGGGGCTT ATTATGCACC TGGTTATACT CAGACCAGTT ACTCCACAGA AGTTCCAAGT ACTTACCGTT 180	г ссябтетете бттббятетя тесесявень вяствтеннь ветбянвене сесететтяя ббббсяббтт 270	360 360 A GARCCCTGG CCCATTATCC TTATGGAGAT GGTAATCGTA GTGTTCCACA ATCACGGCCG 360 M L M U I U V F H N H G R	ICTOTATICORIC CHCARGARAG RIGCGTGGGC TTCTCCTGGT GCTTATGGAR TGGGTGGCCG TTATCCCTGG CCTTCATCAG CGCCCTCAGC 450 L Y D H K K D A H A S P G A Y G M G G R Y P W P S S A P S A	RCCACCCGGC RATCTCTACA TGACTGARAG TACTTCACCA TGGCCTAGCA GTGGCTCTCC CCAGTCACCC CCTTCACCCC CAGTCCAGCA ${ m 540}$ P P G N L Y M T E S T S P W P S S G S P Q S P P S P V Q Q	CCCARGGAT TOTTCATACC COTATAGCCA ATCAGATCAA AGCATGAACC GGCACAACTT TCCTTGCAGT GTCCATCAGT RCGAATCCTC 630 P K D S S Y P Y S Q S D Q S M N R H N F P C S V H Q Y E S S	3 ATTCAGATCT TITGGATTCC CARGICCAGT ATAGTGCTGA GCCTCAGCTG TATGGTAATG CCACCAGTGA 720 S D L L D S Q U Q Y S A E P Q L Y G N A T S D	CATCCCAAC AATCAAGATC AAAGTAGCAG TCTTCCTGAA GAATGTGTAC CTTCAGATGA AAGTACTCCT CCGAGTATTA AAAAAATCAT $f 810$ H $f P$ N $f Q$ $f D$ $f Q$ $f S$ $f S$ $f S$ $f L$ $f P$ $f E$ $f C$ $f U$ $f P$ $f S$ $f D$ $f E$ $f S$ $f T$ $f P$ $f S$ $f I$ $f I$	АСАТОТОСТО СНОЯНОСТЕСТВОЯ ВСИВОВНОТИ СНОВИВНЕНИ СИТИТО СИТИТО В СИТИТО ВОВИТИТО ТНОВИЗАТИТЕ В В В В В В В В В В В В В В В В В В В	АРТОСТАНСС АНСОНЯСТТ ТООЯНСТОСЯ ТТСЯОТТОНЯ АСТОООООСС НООНСТОТ АССОСАООСС НОЯНИЯ В ${\tt Q}$ В	1010
бтянсяссяя	тосстсятяс	сносссянст	сессттсяся	САСЯАСВАВС Н К К D	ARTCTCTACA N L Y M	TCTTCATACC S S Y P	ARCARTGATG N N D D	ARTCARGATC N Q D Q	GAGAAGGTCC E K U Q	NATGCTAACC AAGGAACTTT TG N L T K E L L	ATATTGGAAA I L E
нсоятятсст отянонссяя сяяттесяно	GGGCARATAC TGCCTCATAC TCAGGGGCTT	сятствесяя сявесеяяет сеявтетете	ссяббятятс сбссттсяся бяяссстббя	ACTGTACGAC CACAAGAAAG ATGCGTGGGG	ACCACCCGGC 1	GCCCAAGGAT TCTTCATACC CCTATAGCCA	G G G G G G G G G G G G G G G G G G G	► CCATCCCAAC AATCAAGATC AAAGTAGCAG	ACATGTGCTG H U L I	AATGCTAACC	GATTCAGGCC ATATTGGAAA
							FIG.	4			

689	TTGATGITGC AAGACAAATA TCATTACAGC ACGTTAACTT TTCCATTCGG ATCAAAAAA	TTGATGT
630		2 L D ATAGAGA
54 0 (39)	ST ACTGARATAC CAGAGATCTC ACTITIGATA CIGITITGCA CTICATATGT GCTICTATGT	HarctcGA
6		1 8 L
750	E I Q O E V L 3 r D O H R I D K H I L L L L L L I K Y L L	- Z
360	то втосянятсь выссовтные выстясятсе осстоенной остостенее несеместое	CTGRGAT
0,7	HEIJOHNOON GOCCETTONG MANAMANG TOTTOCTTO TONGCHOLCHTH MANGCCOTCO OMMCOTCETT OOMMACTOOL	HC I GHA
	DEUSXEKNPCIREARRANIEUQTLITYID	О
180	TGGATGAGGT ARGINTIGAA AARAACCCCT GCATCCGGGA AGCCAGGAGA AGAGCAGTGA TCGAGGTGCA AACTCTGATC ACATATATTG	TGGATGA
	EIK NEL LOAQ NPS ELY LSSK TEL QGL IGQL	ш
06	<u> </u>	GAGAAATI

7/39

FIG. 6A

		8/39			
MKVNVSCSSV	QTTIDILEEN	QGEDESILTL	GQLRDRIATD	NDVDVETMKL	50
LHRGKFLQGA	DDVSLSTLNF	KENDKIIVMG	GKNALVDDAG	FKMLMQYEKH	100
NLSNLQKAYD	LNLRDVADLE	RGFLEKPKQV	EMGKKLEKKV	KYFNEEAERH	150
LETLDGMNII	TETTPENQAK	RNREKRKTLV	NGIQTLLNQN	DALLRRLQEY	200
QSVLNGDIPE					210

FIG.6B

		9/39			
ATGCCAGTCG	TGAACATACC	AATCAAAATA	CTTGGTCAGA	ATCAATCACA	50
TAGTCGAAGT	AACTCCTCGT	CTTCTGTTGA	CAACGATCGA	AATCAACCAC	100
CACAGCAGCC	ACCTCAACCG	CAACCACAAC	AGCAATCTCA	GCAACAATAC	150
CAGCAGGCTC	CAAACGTGAA	TACCAATATG	CATCATTCCA	ACGGATTCTC	200
ACCTAACTTC	CCATCTCGTA	GTCCTATTCC	GGACTTTCCC	AGTTTTTCAT	250
CTGGGTTCCC	AAACGATTCT	GAATGGTCTT	CGAATTTCCC	GTCGTTTCCA	300
AATTTCCCAA	GTGGATTCTC	AAATGGAAGT	TCTAATTTCC	CTGATTTTCC	350
AAGATTCGGA	AGAGATGGAG	GACTATCGCC	AAACCCACCG	ATGCAAGGAT	400
ACAGGAGAAG	TCCAACACCA	ACATCAACTC	AATCTCCAAC	TTCTACATTA	450
AGACGCAACT	CTCAGCAGAA	TCAAGCTCCT	CCACAATATT	CTCAGCAACA	500
ACCACAACAA	GCTCAACAAC	GTCAGACAAC	TCCTCCGTCA	ACAAAAGCTT	550
CATCTCGACC	ACCATCTCGT	ACTCGTGAAC	CAAAGGAACC	TGAGGTACCC	600
GAGAGACCAG	CAGTTATTCC	ATTGCCATAT	GAGAAGAAGG	AGAAACCACT	650
GGAGAAGAAA	GGTAGTCGTG	ATTCTGGAAA	GGGTGATGAG	AACCTTGAAG	700
AGAACATTGC	CAAGATCACG	ATCGGAAAGA	ATAATTGCGA	GTTATGTCCG	750
GAACAAGAAA	CGGACGGCGA	CCCATCTCCA	CTAACCTCCC	CAATCACCGA	800
AGGAAAGCCA	AAGAGAGGAA	AGAAACTTCA	ACGTAATCAA	AGTGTTGTTG	850
ATTTCAATGC	CAAGACAATT	GTTACTTTGG	ATAAAATTGA	ATTACAAGTT	900
GAGCAGTTGA	GAAAAAAGC	TGCTGAACTC	GAAATGGAAA	AAGAGCAAAT	950
TCTTCGTTCT	CTAGGAGAAA	TCAGTGTTCA	TAACTGCATG	TTCAAACTGG	1000
AAGAATGTGA	TCGTGAAGAG	ATTGAAGCAA	TCACTGACCG	ATTGACAAAA	1050
AGAACAAAGA	CAGTTCAAGT	TGTTGTCGAA	ACTCCACGAA	ATGAAGAACA	1100
GAAAAAAGCA	CTGGAAGATG	CAACTTTGAT	GATCGATGAA	GTCGGAGAAA	1150
TGATGCATTC	GAATATTGAA	AAGGCTAAGC	TGTGCCTACA	AACCTACATG	1200
AACGCCTGTT	CGTACGAAGA	AACTGCTGGA	GCCACCTGCC	AAAACTTCTT	1250
GAAGATCATA	ATTCAGTGCG	CTGCTGATGA	TCAGAAACGC	ATCAAGCGTC	1300
GTCTGGAAAA	TCTGATGTCT	CAAATTGAGA	ATGCTGAGAG	AACGAAAGCA	1350
GATTTGATGG	ATGATCAAAG	CGAATAG			1377

FIG.7A

		10/39			
MPVVNIPIKI	LGQNQSHSRS	NSSSSVDNDR	NQPPQQPPQP	QPQQQSQQQY	50
QQAPNVNTNM	HHSNGFSPNF	PSRSPIPDFP	SFSSGFPNDS	EWSSNFPSFP	100
NFPSGFSNGS	SNFPDFPRFG	RDGGLSPNPP	MQGYRRSPTP	TSTQSPTSTL	150
RRNSQQNQAP	PQYSQQQPQQ	AQQRQTTPPS	TKASSRPPSR	TREPKEPEVP	200
ERPAVIPLPY	EKKEKPLEKK	GSRDSGKGDE	NLEENIAKIT	IGKNNCELCP	250
EQETDGDPSP	LTSPITEGKP	KRGKKLQRNQ	SVVDFNAKTI	VTLDKIELQV	300
EQLRKKAAEL	EMEKEQILRS	LGEISVHNCM	FKLEECDREE	IEAITDRLTK	350
RTKTVQVVVE	TPRNEEQKKA	LEDATLMIDE	VGEMMHSNIE	KAKLCLQTYM	400
NACSYEETAG	ATCQNFLKII	IQCAADDQKR	IKRRLENLMS	QIENAERTKA	450
DLMDDQSE					458

FIG.7B

		11/39			
ATGTCAGAAA	AGACTAGCAC	AGTTACAATA	CACTATGGAA	ATCAGCGATT	50
TCCGGTAGCA	GTCAATCTAA	ATGAGACGTT	AAGTGAACTG	ATTGATGATT	100
TACTTGAAAC	GACTGAGATT	TCTGAGAAGA	AAGTCAAGCT	TTTTTACGCT	150
GGCAAGCGTT	TAAAAGACAA	AAAAGCCTCG	TTATCAAAAT	TGGGTTTAAA	200
AAATCATAGT	AAAATTCTAT	GTATAAGACC	ACATAAGCAA	CAACGAGGTT	250
CCAAGGAAAA	AGACACGGTT	GAGCCCGCTC	CGAAAGCGGA	AGCGGAGAAT	300
CCTGTATTTT	CGCGTATTTC	TGGAGAAATA	AAAGCCATCG	ATCAGTATGT	350
TGACAAAGAA	CTTTCCCCCA	TGTACGACAA	TTACGTAAAT	AAACCGTCGA	400
ACGATCCAAA	GCAGAAAAAC	AAACAGAAAC	TAATGATAAG	TGAACTACTT	450
TTACAACAGC	TTTTAAAATT	GGATGGAGTT	GACGTACTGG	GCAGCGAGAA	500
ATTGCGTTTT	GAACGGAAGC	AACTTGTTTC	TAAGATCCAA	AAAATGTTGG	550
ATCACGTTGA	CCAAACAAGC	CAAGAAGTGG	CCGCATAG		588

		12/39			
MSEKTSTVTI	HYGNQRFPVA	VNLNETLSEL	IDDLLETTEI	SEKKVKLFYA	50
GKRLKDKKAS	LSKLGLKNHS	KILCIRPHKQ	QRGSKEKDTV	EPAPKAEAEN	100
PVFSRISGEI	KAIDQYVDKE	LSPMYDNYVN	KPSNDPKQKN	KQKLMISELL	150
LQQLLKLDGV	DVLGSEKLRF	ERKQLVSKIQ	KMLDHVDQTS	QEVAA	195

FIG.8B

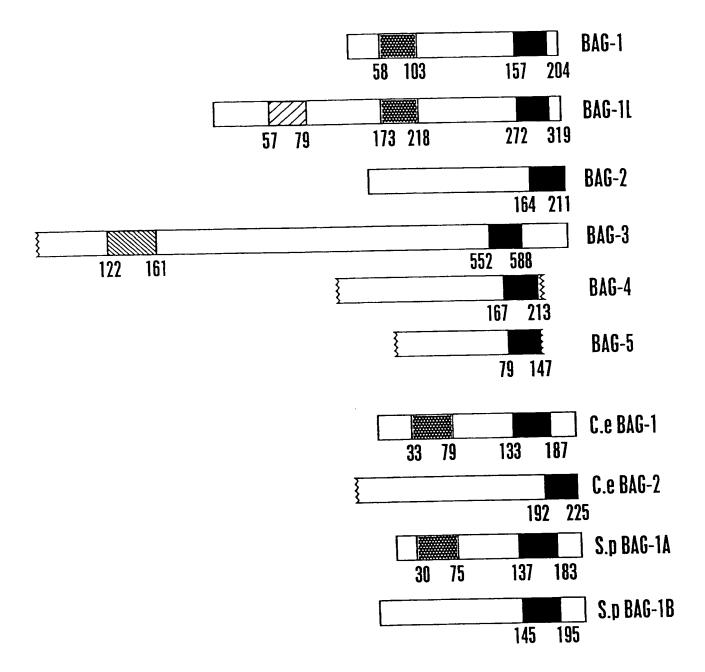
		13/39			
ATGTCTTTT	TTACCCAGTT	GTGTTCTATG	GATAAAAAAT	ATTGGATCTC	50
TCTAGCTGTA	TTGTCAGTTA	CTGTTTTGAT	TAGCGCATTA	TTGAAAAAGA	100
GAGCTACTGA	AACCGAAGAT	ATTGTCGTTG	TTCATTACGA	TGGCGAAAAG	150
TTGAATTTTG	TGTTGCGACA	ACCAAGGCTG	AATATGGTTT	CTTACACTAG	200
TTTTCTTCGT	CGCGTGTGCA	ACGCATTTTC	AGTAATGCCC	GACAAAGCGT	250
CTCTCAAGTT	AAACGGGGTG	ACCCTCAAGG	ATGGTTCACT	TTCCGACCAA	300
AATGTGCAAA	ATGGAAGTGA	ATTAGAGCTC	GAATTACCCA	AACTGAGCCC	350
GGCAATGCAA	CAAATTGAAG	CATATATAGA	TGAGCTTCAA	CAGGATCTCG	400
TCCCTAAAAT	TGAAGCCTTC	TGCCAATCGT	CTCCCGCTTC	GGCACAAGAT	450
GTTCAAGATT	TGCATACACG	CCTTAGTGAA	ACATTGTTGG	CTAGGATGAT	500
AAAATTAGAT	GCTGTTAATG	TTGAAGACGA	CCCAGAAGCT	CGTCTTAAAA	550
GAAAAGAAGC	TATTCGTTTA	TCTCAACAAT	ATTTGAGTAA	ACTAGATTCC	600
ACCAAGAATC	AAAACAAATG	Α			621

FIG. 9A

		14/39			
MSFFTQLCSM	DKKYWISLAV	LSVTVLISAL	LKKRATETED	IVVVHYDGEK	50
LNFVLRQPRL	NMVSYTSFLR	RVCNAFSVMP	DKASLKLNGV	TLKDGSLSDQ	100
NVQNGSELEL	ELPKLSPAMQ	QIEAYIDELQ	QDLVPKIEAF	CQSSPASAQD	150
VQDLHTRLSE	TLLARMIKLD	AVNVEDDPEA	RLKRKEAIRL	SQQYLSKLDS	200
TKNQNK					206

FIG.9B

Fig. 10A

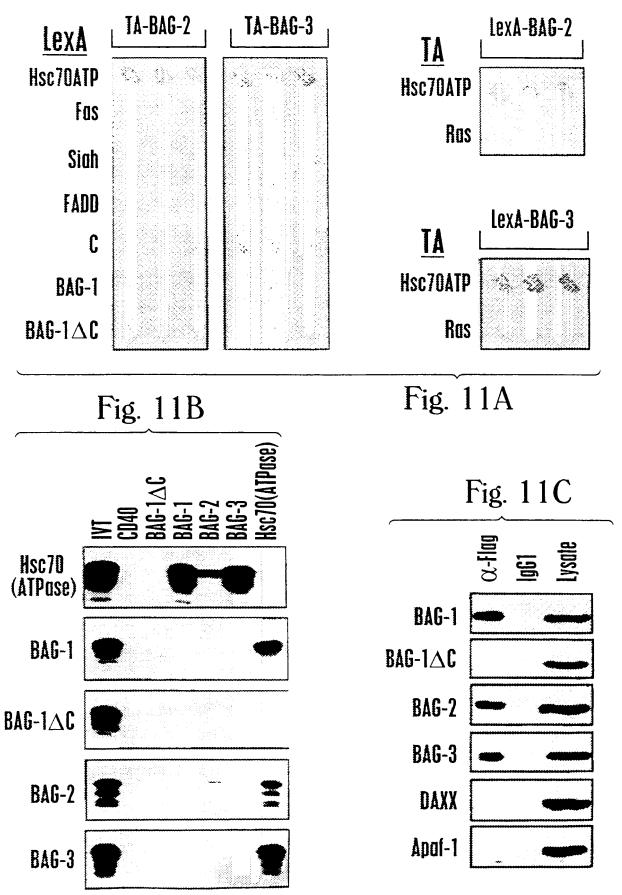


Ubiquitin-Like BAG Domain

WW Nuclear Localization Signal

16/39 SUBSTITUTE SHEET (RULE 26)

Fig. 10B



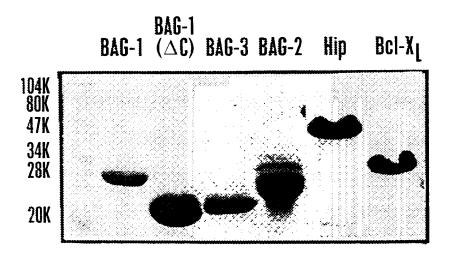
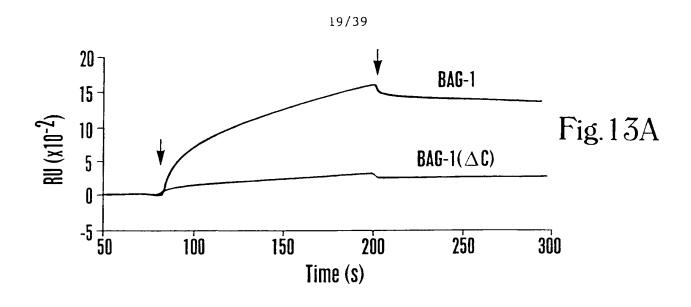
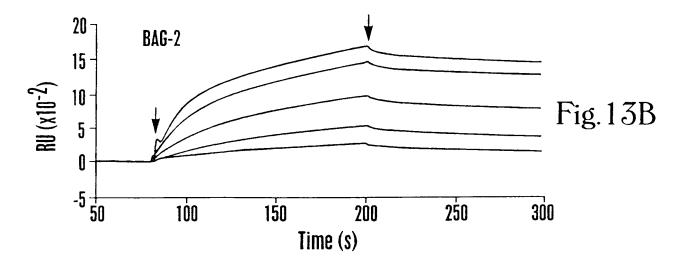
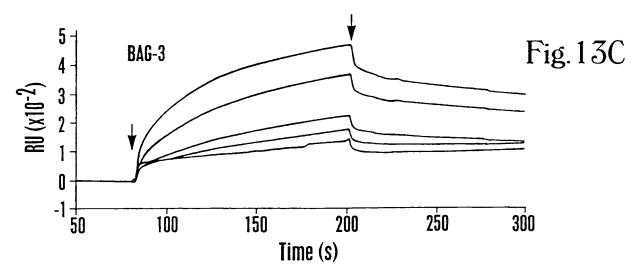
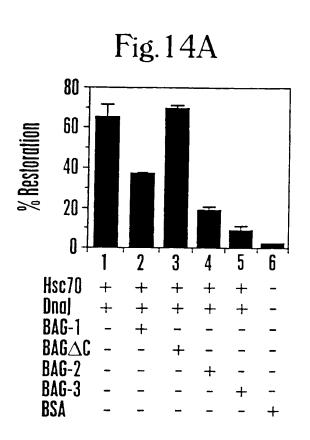


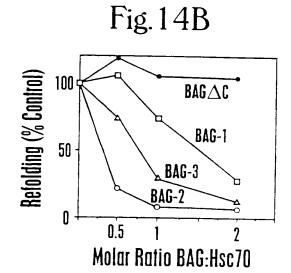
Fig. 12











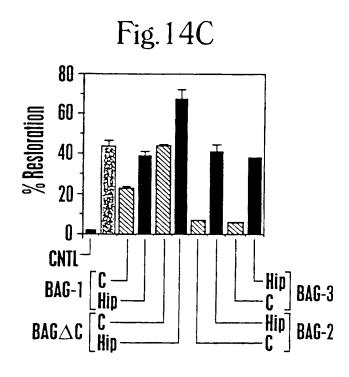


FIG. 15A

GOGGAGCTOC GCATOCAACC COGGGCCGCG GCCAACTTCT CTGGACTGGA 50		•	ACACETCEGO GGCGGAGAGG GGCCCACGGC GGCGGCCCGG CCAGAGACTC 250	GOCAGOATGA GOGOCGOOAO OCACTOGOCO ATGATGOAGG TGGOGTOOGG 350	()	CGCAGACCGG CTGGCCCTTC TTCGTGGACC ACAACAGCCG CACCACTACG 450	TGGAACGACC CGCGCGTGCC CTCTGAGGGC CCCAAGGAGA CTCCATCCTC 500	TGCCAATGGC CCTTCCCGGG AGGGCTCTAG GCTGCCGCCT GCTAGGGAAG 550	GCCACCCTGT GTACCCCCAG CTCCGACCAG GCTACATTCC CATTCCTGTG 600	CTCCATGAAG GCGCTGAGAA CCGGCAGGTG CACCCTTTCC ATGTCTATCC 650	CCAGCCTGGG ATGCAGCGAT TCCGAACTGA GGCGGCAGCA GCGGCTCCTC 700	AGAGGTCCCA GTCACCTCTG CGGGGCATGC CAGAAACCAC TCAGCCAGAT 750	AAACAGTGTG GACAGGTGGC AGCGGCGGCG GCAGCCCAGC CCCCAGCCTC 800	CCACGGACCT GAGCGGTCCC AGTCTCCAGC TGCTCTGAC TGCTCATCCT 850	CATOCTOCTO GGCCAGOCTG COTTOCTCOG GCAGGAGCAG CCTGGGCAGT 900	CACCAGCTCC CGCGGGGGTA CATCTCCATT CCGGTGATAC ACGAGCAGAA 950	CGTTACCCGG CCAGCAGCCC AGCCCTCCTT CCACAAAGCC CAGAAGACGC 1000	ACTACCCAGC GCAGAGGGGT GAGTACCAGA CCCACCAGCC TGTGTACCAC 1050	AAGATCCAGG GGGATGACTG GGAGCCCCGG CCCCTGCGGG CGGCATCCCC 1100	attcaggtca tctgtccagg gtgcatcgag ccgggagggc tcaccagcca 1150	GGAGCAGCAC GCCACTCCAC TCCCCTCGC CCATCCGTGT GCACACCGTG 1200	GTCGACAGGC CTCAGCAGCC CATGACCCAT CGAGAAACTG CACCTGTTTC 1250	CCAGCCTGAA AACAAACCAG AAAGTAAGCC AGGCCCAGTT GGACCAGAAC 1300	
GCGGAGCTCC GCATCCAA(CCTCTGGCAG CGAGGAGG	ACGTCACCCC CGCCTTTA	ACACGTCGGC GGCGGAGAG	CCCAGCATGA GCGCCGCC	CAACGGTGAC CGCGACCC	CGCAGACCGG CTGGCCCT	TGGAACGACC CGCGCGTG(TGCCAATGGC CCTTCCCGG	GCCACCCTGT GTACCCCC	CTCCATGAAG GCGCTGAG	CCAGCCTGGG ATGCAGCG/	AGAGGTCCCA GTCACCTC	AAACAGTGTG GACAGGTG(CCACGGACCT GAGCGGTCC	CATCCTCCTC GGCCAGCCT	CACCAGCTCC CGCGGGGG	CGTTACCCGG CCAGCAGCC	ACTACCCAGC GCAGAGGGC	AAGATCCAGG GGGATGACT	GTTCAGGTCA TCTGTCCAG	GGAGCAGCAC GCCACTCCA	GTCGACAGGC CTCAGCAG(CCAGCCTGAA AACAAACCA	CC+C <c cc+cc+ccc+<="" td=""></c>

FIG.15B

FIG. 15C

550 200 සි 200 320 VPGQVQVYEL QPSNLEADQP LQAIMEMGAV AADKGKKNAG NAEDPHTETQ EGAENRQVHP FHVYPQPGMQ RFRTEAAAAA PQRSQSPLRG MPETTQPDKQ CGQVAAAAA QPPASHGPER SQSPAASDCS SSSSSASLPS SGRSSLGSHQ APAEATPPKP GEAEAPPKHP GVLKVEAILE KVQGLEQAVD NFEGKKTDKK **QGDDWEPRPL RAASPFRSSV QGASSREGSP** ARSSTPLHSP SPIRVHTVVD VSQKPPPPSE KVEVKVPPAP VPCPPPSPGP SAVPSSPKSV ATEERAAPST YLMIEEYLTK ELLALDSVDP EGRADVROAR RDGVRKVOTI LEKLEGKAID RPQQPMTHRE TAPVSQPENK PESKPGPVGP ELPPGHIPIQ VIRKEVDSKP DPRVPSEGPK ETPSSANGPS REGSRLPPAR EGHPVYPOLR PGYIPIPVLH PRGYISIPV IHEONVTRPA AQPSFHKAQK THYPAQRGEY QTHQPVYHKI APEATAAATS NPSSMTDTPG NPAAP

50

MSAATHSPMM QVASGNGDRD PLPPGWEIKI DPQTGWPFFV DHNSRTTTWN

								24/39						
90	1.80	270	390	450	240	P30	720	970	900	940	1080	1170	1260	1350
TCCCTTTATC	TCATAAAGGT	GCCAGCGCC	CAACGGTGAC N G D	CACCACTACG T T T	6CTGCCGCCT L P P	CCGCCAGGTG R Q V	CTCACCTCTG S P L	CCACGGACCT H G P	CCTGGGCAGT L G S	CCACAAAGCC H K A	GGAGCCCGG E P R	GCCACTCCAC P L H	CCAGCCTGAA Q P E	GGCCCAGTT GGACCAGAAC TCCCTCCGG ACACATCCA ATTCAAGTGA TCCGCAAAGA GGTGGATTCT 1355 5 P V G P E L P P G H I P I Q V I R K E V D S
AGTTGCTACC	CGCCTTTAAT	66(6(((664	7666676666 A S 6	ACAACAGCCG N S R	AGGGCTCTAG G S R	GCGCTGAGAA A E N	AGAGGTCCCA R S Q	CCCCAGCCTC P A S	GCAGGAGCAG R S S	AGCCCTCCTT PSF	GGGATGACTG D D W	GGAGCAGCAC S S T	CACCTGTTTC P V S	TCCGCAAAGA R K E
CTAGCCGGCC	ACGTCACCCC	CCAGAGACTC	ATGATGCAGG M M Q V	TTCGTGGACC F V D H	CCTTCCCGGG P S R E	CTCCATGAAG L H E G	GCGGCTCCTC A A P A	GCAGCCCAGC A A A P	CCTTCCTCCG P S S G	CCAGCAGCCC P A A A	AAGATCCAGG K I Q G	TCACCAGCCA S P A R	CGAGAAACTG R E T A	ATTCAAGTGA I Q V I
CCAGAAGTTT	CTCTCTGGCC	99))))99)99	CCACTCGCCC H S P	CTGGCCCTTC WPF	TGCCAATGGC A N G	CATTCCTGTG I P V	66C66CA6CA A A A	AGCGGCGGCG A A A	GGCCAGCCTG A S L	CGTTACCCGG V T R	TGTGTACCAC V Y H	CCGGGAGGGC R E G	CATGACCCAT M T H	ACACATCCCA H I P
CTGGACTGGA	GGCT ATTTCCAGAC ACTTCCACCC	CGCCCACGGC	GCGCCGCCAC A A T	CGCAGACCGG	CTCCATCCTC P S S	GCTACATTCC Y I P	TCCGAACTGA R T E	GACAGGTGGC Q V A	CATCCTCCTC S S S	ACGAGCAGAA E Q N	CCCACCAGCC H Q P	GTGCATCGAG A S S	CTCAGCAGCC a a P	TCCCTCCT66
CGCG GCCAACTTCT	ATTTCCAGAC	6606646466	CCCAGCATGA N S	AAGATCGACC K I D P	CCCAAGGAGA P K E T	CTCCGACCAG L R P G	ATGCAGCGAT M & R F	AAACAGTGTG K Q C G	TGCTCATCCT C S S S	CCGGTGATAC P V I H	GAGTACCAGA E Y Q T	TCTGTCCAGG S V Q G	GTCGACAGGC V D R P	GGACCAGAAC G P E L
9)9))999))	CGAGGAGGCT	ACACGTCGGC	CAGACCCCAA	ATGGGAGATC W E I	CTCTGAGGGC S E G	GTACCCCCAG Y P Q	CCAGCCTGGG Q P G	TCAGCCAGAT	TGCCTCTGAC A S D	CATCTCCATT I S I	GCAGAGGGGT Q R G	GTTCAGGTCA F R S	GCACACCGTG H T V	AGGCCCAGTT G P V
GCATCCAACC	CCTCTGGCAG	99)))1100	CCCCAGCGG	16((((((6	CGCGCGTGCC R V P	GCCACCCTGT W P W	ATGTCTATCC V Y P	CAGAAACCAC E T T	AGTCTCCAGC S P A	CGCGGGGTA R G Y	ACTACCCAGC Y P A	CGGCATCCCC A S P	CCATCCGTGT I R V	AAAGTAAGCC S K P
GCGGAGCTCC	1001001))9)99)))9	CGCACCCGCG CCCCAGCGGG CAGACCCC	CGCGACCCTT R D P L	TGGAACGACC CGCGCGTGCC CTCTGAGG	GCTAGGGAAG A R E G	CACCCTTTCC H P F H	CGGGGCATGC R G M P	GAGCGGTCCC E R S A	CACCAGCTCC H & L P	CAGAAGACGC Q K T H	CCCCTGCGGG P L R A	TCCCCTCGC CCATCCGTGT GCACACCG	AACAAACCAG N K P E
							(J)	1						

1710 1620 1440 1530 TTCCCCCAAG AGTGTGGCTA CAGAAGAGAG GGCAGCCCCC AGCACTGCCC CTGCAGAAGC TACACCTCCA GTAGACAACT TTGAAGGCAA GAAGACTGAC AAAAAGTACC TGATGATCGA AGAGTATTTG ACCAAAGAGC TGCTGGCCCT GGATTCAGTG AAACCTGTTT CCCAGAAGCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA AGTTCCCCCT GCTCCAGTTC CTTGTCCTCC TCCCAGCCCT AAACCAGGAG AAGCCGAGGC TCCCCCAAAA CATCCAGGAG TGCTGAAAGT GGAAGCCATC CTGGAGAAGG TGCAGGGGCT GGAGCAGGCT __ 9 STAP ~ ~ ~ ~ ۵. <u>а</u> **~** ... ⊢--H P G V L K V ∞ ш ш _ v ∧ K K ---_ _ _ ب م س a. (161((((1 п 6 7 о О > 9600011016

1,800 TGTGCGTCAG GCCAGGAGAG ACGGTGTCAG GAAGGTTCAG ACCATCTTGG AAAAACTTGA ACAGAAAGCC **∀ ∀** K V Q T I L E K L E ک ح 9 <u>~</u> o∠ •× У В В GACCCCGAGG GACGAGCCGA A 0 <u>م</u>

ATTGATGTCC CAGGTCAAGT CCAGGTCTAT GAACTCCAGC CCAGCAACCT TGAAGCAGAT CAGCCACTGC AGGCAATCAT GGAGATGGGT 0 P L 0 0 V 3 z v - 7 - 1 - 3 7 ~ 7 7 9

CAAGAAAAAT GCTGGAAATG CAGAAGATCC CCACACAGAA ACCCAGCAGC CAGAAGCCAC AGCAGCAGCG

H T E T Q Q P

9 ¥

GCCGTGGCAG CAGACAAGGG

2070 ACTTCAAACC CCAGCAGCAT GACAGACACC CCTGGTAACC CAGCAGCACC CTAGCCTCTG CCCTGTAAAA ATCAGACTCG GAACCGATGT **9**

2430 2250 2340 TTAAGTCAGT TGGTTTTTAT TAGCTGCTTG GTATGCAGTA ACTTGGGTGG GATGGGGAGT CAATTACCCA TCACATAAAT ATGAAACATT TATCAGAAAT GTTGCCATTT TAATGAGATG CTTGTTCTGC AGCCCTGTCT ACTTGGGCAC CCCCACCACC TGTTAGCTGT GGTTGTGCAC GGCTAAAAAG GAAAATGATG CTTTTCTTCT ATATTCTTAC TCTGTACAAA TAAAGAAGTT TCTCATAATT AAAATACCTG ACTTTAGAGA GAGTAAAATG TGCCAGGAGC CATAGGAATA TCTGTATGTT TTGCATGCAT TTCAGAGACT GAATTTTAAG CTAATAAAG GAGAAGTTT AACCCCGTTG **IGGAGGGGTA** AGGCAAAACA

FIG. 16A

CATACA CA CARA CARATICAGA GCAGCAGATO CCATATOGGO	20
	8
	150
	200
	250
	300
	350
	400
	450
	200
	550
	009
ATOTAL APOTAGGACTG TOAGACTGAA GCACCCCTC TTAGGGGCA	029
ALCIAL COO MACO CONTINUE CACAGAACCO TGGAATGACC CTGCCCCATT	700
ATCOTTATES AGATESTAST CSTAGTETTC CACAATCAGG ACCGACTGTA	750
CANCACA A A GATGCGTG GCCTTCTCCT GGTGCTTATG GAATGGGTGG	800
	850
ACATE ACTE A AGTACTTCA CCATGGCCTA GCAGTGGCTC TCCCCAGTCA	006
	950
CCAATCAGAT CAAAGCATGA ACCGGCACAA CTTTCCTTGC AGTGTCCATC	1000
AGTACGAATC CTCGGGGGACA GTGATCAATG AAGATTCAGA TCTTTTGGAT	1050
TOCCA GETOD AGTATAGTGC TGAGCCTCAG CTGTATGGTA ATGCCACCAG	1100
TOACCATCCC AAGAATCAAG ATCAAAGTAG CAGTCTTCCT GAAGAATGTG	1150
TACCTTCAGA TGAAAGTACT CCTCCGAGTA TTAAAAAAAT CATACATGTG	1200
CTOBABABB TOCAGTATCT TGAACAAGAA GTAGAAGAAT TTGTAGGAAA	1250
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1300
AAAGACACACACACACACACACACACACACACACACACA	

PCT/US99/21053

27/39

FIG. 16B

TTTTEGAACT GGATTCAGTT GAAACTGGGG GCCAGGACTC TGTACGGCAG 1350	1350
OCCAPACA A A B A B B C TETT B TA B A TTCA B G C C A TA C T B C A A A A A A A A A A A A A A A A A A	1400
SCCAGARANA AND A ARGE ATTTAGA A A GTGGA A GTGGA A GTGGA A GTGGA A GTGA A A GTGGA A GTGA A A GTGGA A GTGA A A A	1450
AAAAAAGGA IIAIGAAAGGTTAAT TACCCTCTTT TTGAAATGCC	1500
TOTICACCARA GARAGGAT ACATTCCAGC TITTCCTTTG ATTITATACT	1550
TO A A A A A A CATE OCA BAGGAAT GGAAGAATAT TTTAGTCATG AAGTTGTTTT	1600
CASTITION OF A TIGA A T	1650
CAGILLICACA COLOCTA AAAAATTAT GGATATCTAC AAGCTGCTTA	001
CCAAGINGACIONO CAGGAAACA CAGTICACAC AACAGGCTTA TCAGAAACCT	1750
HACCAGCAG GAGGGAAAA ATTTBAGACA AACAGGATGT GTTTTTTAA	0081
ACCAGA GAA ACTGGATATA ATTTACAGAGAGAGAGAGAGAGATAGAAGAATAGAAGAATAGAAGA	1850
ACATCTGGAT ATCTTGTCAC ALLITIGIAC ALLGIGACIG CLITOCACAL	
ATACTTICATG TGTAATTATA GCTTAGACTT TAGCCTTCTT GGACIICIGI	0 1
THE TITIEST TATTICE AGT TTACAAATAT AGTATTATTC TCTAAAAAA	0 5 b/
AAAAAAAA AAAAA	9961

28/39

FIG. 16C

DAWASPGAYGMGGRYPWPSSAPSAPPGNLYMTESTSPWPSSGSPQSPPSPPVQQPKDSSYPYSQSDQSMNRHNFPCSVHQ EPGRAGGSHQEQPPYPSYNSNYWNSTARSRAPYPSTYPVRPELQGQSLNSYTNGAYGPTYPPGPGANTASYSGAYYAPGY MSALBRSGYGPSDGPSYGRYYGPGGGDVPVHPPPPLYPLRPEPPQPPISWRVRGGGPAETTWLGEGGGGGGGYYPSGGAWP TQTSYSTEVPSTYRSSGNSPTPVSRWIYPQQDCQTEAPPLAGQVPGYPPSQNPGMTLPHYPYGDGNRSVPQSGPTVRPQE YESSGTVINEDSDLLDSQVQYSAEPQLYGNATSDHPNNQDQSSSLPEECVPSDESTPPSIKKIIHVLEKVQYLEQEVEEF **VGKKTDKAYWLLEEMLTKELLELDSVETG**GQDSVRQARKEAVCKIQAILEKLEKK**G**L

B	140	270	360	n 50	240	P 30	720	970
cggtcgfcc	TCCCCAGCCT	TCCCTCGGGA	GAATTCTACT	TGGAGCGTAT	CAGTTACTCC	TCAGACTGAA	AGATGGTAAT	CCGTTATCCC
g p S	P Q P	P S G	N S T	G A Y	S Y S	Q T E	D G N	R Y P
GCCCCAGTGA	GCCCTGAACC	ATGGCTACTA	CTAACTATTG	CTTATACAAA	ATACTCAGAC	AGCAGGACTG	ATCCTTATGG	GAATGGGTGG
P S D	P E P	G Y Y	N Y W	Y T N	T Q T	A D C	P Y G	M G G
CGGTGGGAGC GGGGCGGGAA GCGCTTCAGG GCAGCGGATC CCATGTCGGC CCTGAGGCGC TCGGGCTACG GCCCCAGTGA CGGTCCGTCC	TACGCCGCT ACTACGGGCC TGGGGGTGGA GATGTGCCGG TACACCCACC TCCACCCTTA TATCCTCTTC GCCCTGAACC TCCCCAGCCT	CCCATTTCCT GGCGGGGGGGGGGGGGGGGGGGGGGGGGG	GGCGCCTGGC CAGAGCCTGG TCGAGCCGGA GGAAGCCACC AGGAGCAGCC ACCATATCCT AGCTACAATT CTAACTATTG GAATTCTACT	GCGAGATCTA GGGCTCCTTA CCCAAGTACA TATCCTGTAA GACCAGAATT GCAAGGCCAG AGTTTGAATT CTTATACAAA TGGAGCGTAT	GGTCCAACAT ACCCCCCAGG CCCTGGGGCA AATACTGCCT CATACTCAGG GGCTTATTAT GCACCTGGTT ATACTCAGAC CAGTTACTCC	ACAGAAGTTC CAAGTACTTA CCGTTCATCT GGCAACAGCC CAACTCCAGT CTCTCGTTGG ATCTATCCCC AGCAGGACTG TCAGACTGAA	GCACCCCCTC TTAGGGGGCA GGTTCCAGGA TATCCGCCTT CACAGAACCC TGGAATGACC CTGCCCCATT ATCCTTATGG AGATGGTAAT	CGTAGTGTTC CACAATCAGG ACCGACTGTA CGACCACAAG AAGATGCGTG GGCTTCTCCT GGTGCTTATG GAATGGGTGG CCGTTATCCC
M S A L R R S G Y G P S D G P S	Y G R Y Y G P G G G D V P V H P P P P L Y P L R P E P Q P		G A W P E P G R A G G S H Q E Q P P P S Y N S N Y W N S T	A R S R A P Y P S T Y P V R P E L Q G Q S L N S Y T N G A Y	G P T Y P P G P G A N T A S Y S G A Y Y A P G Y T Q T S Y S	T E V P S T Y R S S G N S P T P V S R W I Y P Q Q D C Q T E	A P P L R G Q V P G Y P P S Q N P G M T L P H Y P Y G D G N	RSVPQSSVPGSGPGACGACTGTA CGACCACAAG AAGATGCGTG GGCTTCTCCT GGTGCTTATG GAATGGGTGG CCGTTATCCC
CCTGAGGCGC	TCCACCCTTA	GGGAGAAGGC	ACCATATCCT	GCAAGGCCAG	GGCTTATTAT	CTCTCGTTGG	TGGAATGACC	GGCTTCTCCT
L R R S	P P L	G E G	P Y P	a G a	A Y Y	S R W	G M T	A S P
CCATGTCGGC	TACACCCACC	CCACCTGGCT	AGGAGCAGCC	GACCAGAATT	CATACTCAGG	CAACTCCAGT	CACAGAACCC	AAGATGCGTG
M S A	H P P	T W L	E Q P	P E L	YSG	T P V	Q N P	D A W
GCAGCGGATC	GATGTGCCGG	CCGGCGGAGA	GGAAGCCACC	TATCCTGTAA	AATACTGCCT	GGCAACAGCC	TATCCGCCTT	CGACCACAAG
	D V P V	P A E T	G S H A	Y P V R	N T A S	G N S P	Y P P S	R P Q E
GCGCTTCAGG	TGGGGGTGGA	9 9 9	TCGAGCCGGA	CCCAAGTACA	CCCTGGGCA	CCGTTCATCT	GGTTCCAGGA	ACCGACTGTA
	G G G)99)9999)	R A G	PST	P G A	R S S	V P G	P T V
6666C6G6AA	ACTACGGGCC Y G P	GGCGGGTGCG R V R	CAGAGCCTGG E P G	GGGCTCCTTA A P Y	ACCCCCA66	CAAGTACTTA S T Y	TTAGGGGGCA R G A	CACAATCAGG Q S G
CGGTGGGAGC	TACGGCCGCT Y G R Y	CCCATTTCCT P I S W	9 N P 9	GCGAGATCTA A R S R	GGTCCAACAT G P T Y	ACAGAAGTTC T E V P	GCACCCCTC A P P L	CGTAGTGTTC R S V P

900	940	1080	1170	1260	1,350	1440	1530	1,620	1710	1,800	1,690	1980
TCCCCAGTCA P Q S	CTTTCCTTGC F P C	TGAGCCTCAG E P A	TGAAAGTACT E S T	AAAGACAGAC K T D	TGTACGGCAG V R Q	AAGTGGAAGC	ACATTCCAGC	TTTAGTCATG AAGTTGTTTT CAGTTTTCAG ACGAATGAATG	TTACCAGCAG	GTTTTTTAA	TAGCCTTCTT	
GCAGTGGCTC S G S	ACCGGCACAA R H N	AGTATAGTGC Y S A	TACCTTCAGA PSD	TTGTAGGAAA V G K	GCCAGGACTC Q D S	ATTTAGAACA	AAGAAGCAAT	CAGTTTTCAG	AAAAATTTAT GGATATCTAC AAGCTGCTTA TTACCAGCAG	AACAGGATGT	GCTTAGACTT	AAAAA
CCATGGCCTA P W P S	CAAAGCATGA Q S M N	TCCCAAGTCC S a V a	GAAGAATGTG E E C V	GTAGAAGAAT V E E F	GAAACTGGGG E T G G	TTATGAAAGG L	TGTTGATGAC	AAGTTGTTTT	GGATATCTAC	ATTTGAGACA	TGTAATTATA	AAAAAAAA
AAGTACTTCA S T S	CCAATCAGAT Q S D	TCTTTTGGAT L L D	CAGTCTTCCT S L P	TGAACAAGAA E Q E	GGATTCAGTT D S V	AAAAAAAGGA K K G	TTGAAATGCC	TTTAGTCATG	AAAATTTAT	ACTGGATATA	ATACTTCATG	TCTAAAAAA
ACATGACTGA M T E	ACCCCTATAG P Y S	AAGATTCAGA D S D	ATCAAAGTAG Q S S	TCCAGTATCT Q Y L	TTTTGGAACT L E L	AAAAATTAGA K L E	TACCCTCTTT	GGAAGAATAT	TCACTCCTTA	ACCAGATGAA	CTTTCAACAT	AGTATTATTC
GGCAATCTCT G N L Y	GATTCTTCAT	GTGATCAATG V I N E	AACAATCAAG N A A D	CTGGAGAAGG L E K V	ACCAAGGAAC T K E L	GCCATACTGG A I L E	TTAGGTTAAT	GCAAAGGAAT	CCAAGTAGAC	TCAGAAACCT	ATTGTGACTG	TTACAAATAT
AGCACCACCC A P P	GCAGCCCAAG A P K	CTCGGGGACA S G T	TGACCATCCC D H P	CATACATGTG I H V	AGAAATGCTA E M L	TAAGATTCAG K I Q	GAACACTTGA	TGAAAAACTG	ACCAATATTG	AACAGGCTTA	ATTTTTGTAC	TATTTGCAGT
TGGCCTTCAT CAGCGCCCTC AGCACCACCC GGCAATCTCT ACATGACTGA AAGTACTTCA CCATGGCCTA GCAGTGGCTC TCCCCAGTCA W P S S A P S A P P G N L Y M T E S T S P W P S S G S P Q S	CCCCAGTCCA P V Q	AGTGTCCATC AGTACGAATC CTCGGGGACA GTGATCAATG AAGATTCAGA TCTTTTGGAT TCCCAAGTCC AGTATAGTGC TGAGCCTCAG SVH QYESSGTVINEDSDLLDSQVQYQ	CTGTATGGTA ATGCCACCACCATCCC AACAATCAAG ATCAAAGTAG CAGTCTTCCT GAAGAATGTG TACCTTCAGA TGAAAGTACT L Y G N A T S D H P N N Q D Q S S L P E E C V P S D E S T	TTAAAAAAT K K I	GGCTTCTGGA L L E	GCCAGAAAAG AGGCTGTTTG TAAGATTCAG GCCATACTGG AAAAATTAGA AAAAAAGGA TTATGAAAGG ATTTAGAACA AAGTGGAAGC A R K E A V C K I Q A I L E K L E K K G L	CTGTTACTAA CTTGACCAAA GAACACTTGA TTAGGTTAAT TACCCTCTTT TTGAATGCC TGTTGATGAC AAGAAGCAAT ACATTCCAGC	ITTTCCTTTG ATTTTATACT TGAAAACTG GCAAAGGAAT GGAAGAATAT	CTATGGAGTT	CACTTCACAC	ATCTTGTCAC	GGACTTCTGT TITGTTTTGT TATTTGCAGT TTACAAATAT AGTATTATTC TCTAAAAAA AAAAAAAAA AAAAA
TGGCCTTCAT W P S S	CCCCCTTCAC P P S P	AGTGTCCATC S V H Q	CTGTATGGTA L Y G N	CCTCCGAGTA P P S I	AAAGCATACT K A Y W	GCCAGAAAG A R K E	CTGTTACTAA	1111((1116	TAATAGGAAA	GAGGGAAACA	ACATCTGGAT	GGACTTCTGT

FIG. 17A

OCCOCCCC COCCCCCC CONGAAGACG COCGGAGCGG CIGCIGCAGC	2
CAGTAGCGGC COCTTCACCG GCTGCCCCGC TCAGACCTAG TCGGGAGGGG	8
TECCAGGCAT GCAGCTGGGG GCCCAGCTCC GGTGCCGCAC CCCGTAAAGG	150
GCTGATCTTC CACCTCGCCA CCTCAGCCAC GGGACGCCAA GACCGCATCC	200
AATTCAGACT TCTTTTGGTG CTTGTGAAAC TGAACACAAC AAAAGTATGG	250
ATATGGGAAA CCAACATCCT TCTATTAGTA GGCTTCAGGA AATCCAAAAG	300
GAAGTAAAAA GTGTAGAACA GCAAGTTATC GGCTTCAGTG GTCTGTCAGA	350
TGACAAGAAT TACAAGAAC TGGAGAGAAT TCTAACAAAA CAGCTTTTTG	400
AAATAGACTO TGTAGATACT GAAGGAAAAG GAGATATTCA GCAAGCTAGG	450
AAGCGGGCAG CACAGGAGAC AGAACGTCTT CTCAAAGAGT TGGAGCAGAA	200
TGCAAACCAC CCACACGGGA TTGAAATACA GAACATTTTT GAGGAAGCCC	550
AGTOCOTOGT GAGAGAGAA ATTGTGCCAT TITATAATGG AGGCAACTGC	009
GTAACTGATG AGTTTGAAGA AGGCATCCAA GATATCATTC TGAGGCTGAC	650
ACATGITAAA ACTGGAGGAA AAATCTCCTT GCGGAAAGCA AGGTATCACA	200
CTTTAACCAA AATCTGTGCG GTGCAAGAGA TAATCGAAGA CTGCATGAAA	750
AAGCAGCCTT CCCTGCCGCT TTCCGAGGAT GCACATCCTT CCGTTGCCAA	800
AATCAACTTC GTGATGTGTG AGGTGAACAA GGCCCGAGGG GTCCTGATTG	850
CACTTCTGAT GGGTGTGAAC AACAATGAGA CCTGCAGGCA CTTATCCTGT	006
GTGCTCTCGG GGCTGATCGC TGACCTGGAT GCTCTAGATG TGTGCGGCCG	950
GACAGAAATC AGAAATTATC GGAGGGAGGT AGTAGAAGAT ATCAACAAAT	1000
TATTGAAATA TCTGGATTTG GAAGAGGAAG CAGACACAAC TAAAGCATTT	1050
GACCTGAGAC AGAATCATTC CATTTTAAAA ATAGAAAAGG TCCTCAAGAG	1100
AATGAGAGAA ATAAAAAATG AACTTCTCCA AGCACAAAAC CCTTCTGAAT	1150
TGTACCTGAG CTCCAAAACA GAATTGCAGG GTTTAATTGG ACAGTTGGAT	1200
GAGGTAAGTC TTGAAAAAA CCCCTGCATC CGGGAAGCCA GGAGAAGAGC	1250
AGTGATCGAG GTGCAAACTC TGATCACATA TATTGACTTG AAGGAGGCCC	1300

FIG. 17B

THEASAAAAA AAAAGTATTT GCTTGTGAGG AGCACCCATC CCATAAAGCC	1350
STOTES A A CRITICISA A CHICTORAG ATCCAGGGAG AAGTTCTTTC	1400
ATTRATER ANTOGRACOG ATAGGACTA CATCOGGOTG GAAGAGCTGC	1450
TOACCAAGCA GCTGCTAGCC CTGGATGCTG TTGATCCGCA GGGAGAAGAG	1500
A A GTG TGCCAGGAA ACAAGCTGTG AGGCTTGCGC AGAATATTCT	1550
CABOTATOTO GACOTGAAT CTGATGAATG GGAGTACTGA AATACCAGAG	1600
	1650
	1700
TGATTGAAGC AAATTCTATT CAGTATCTGC TGCTTTTGAT GTTGCAAGAC	1750
AAATATCATT ACAGCACGTT AACTTTTCCA TTCGGATCAT TATCTGTATG	1800
-	1850
AAACAAAAT GAGGCAGCTT TTGTAGATTT TAAATGGGTT GTGCAAGCAT	1900
TA A A A TOCAGE GICTITICAGE ATCTAGAACT AGGCATAACC TTACATAATA	1950
CTAGGAAAAT TATGAGAAAG GGGAAATTTT TGGTTAAATA AGAGTAAGGT	2000
TCAAACACAA GCAGTACATG TTCTGTTTCA TTATGCTCGA TAGAAGGCTT	2050
TITITICACT TATAAGGCCT GATTGGTCCT ACCCAGCTTA ACGGGGTGGG	2100
	2150
GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2200
TRETTITIET GIGCACATA AGAAAATTAT GAAAACTAAT AGCCAAAAA	2250
COTTIGAGAT TECATTAAAG AGAAGGGATA AAGGACCAGC AATAATACCT	2300
TETAAGTTGC TTTTGTTGT AAAATCTGAG CTTATAGTTT TCCTTAGTGA	2350
GTAAATTCAT AAGGATGGGA ACATTTAAAT TAAGTTAATG GGCCTTTAAA	2400
AAAAAAAAA GAAACACTCA TACCTGTAGT TGGAGGATGA ATACTGGAGA	2450
CERETTACCA ATGLOGGTT ATACTAAAAC TAAATCAGAA AGTCTGAATG	2500
TARCACATAA TGGTTCTCTT CTGTTGTCCA AGGCTGTAAA ATGGACAGCC	2550
TTGTCACACC TCCCCGGTGC TGTTTTACAA CGTGAGGGTA GACGCTGTCA	2600

FIG. 17C

GTAACCCAGA GGGACCAGGC CTTCCTAGGT TITCTAGGCA GTCAGCTGTT	2650
AACCACTCAC TTAGTAAATG TCATAACTAC ACCTGCTCCA GGACCAATCA	2700
GTGAAACCTG CTCGGAATTA AAGGCTTCCT CTGGGTGCCT GCTGAACAAC	2750
TGAGCTCATG TCATGGGCAT GTGGTGGTTT CTCTGTTGCC TGAAAGAGCC	2800
ATTAAAGTCA GTCGTGCGTG AAGCATCTCT CTTCTAAAGG ATGTGTATTT	2850
CCATAAATGC TITCTGAGGA TCCGGTACAA AATGATTTCC CAAAGTTCTG	2900
AAGTGCCTTG AGAACATGTG GGTCCGAGTG TTATAACAGA CTCCTCCCCC	2950
GGGTCACCTT TTGCCTGGTC ATCCTGTTAG AGTACATCTT TGGAAATCCA	3000
GGGTAATATT CTCTTTCAGA GATGCTCATT GTGTAACTCT GTGTAGGGAG	3050
ATAGTCACTT TAAACAGCTC AAAGTAGCTA GCTAAAGGAG TAGCCTTAAA	3100
TACCTAAAAG ATGACAGAAG CATAGCCCTT AACAAATCTT CAGCTTGTCT	3150
CTCAGTATTT CCCAATCATG AAAATCCCTT GCTATGTCTT TCCTACTAGA	3200
AATGTTCTAG AATCGCTGGA CGGTGGGGTC AGAGGGCAGT CGGTATTTAG	3250
GCCGTGAGCT TCCCATACTA CTGCAGGTCC AACTCCTGGC AACCGCGGGC	3300
TCAAGGCAGG TCATTGGAAT CCACGTTTTG GCCACAGTAG TTGTAGGATT	3350
GCTTTTCTGT ATCATATTT TAGAATGCTC TTAAAATCTT GAGGAAGAGT	3400
TITIVATITITI TATITIATITI TGAGATGGAG TCTCTGTTGC CCAGGCTGCA	3450
GTGCAGTGGT GCCATCTCAG CTCACTGCAA CCTCCACCTC CCAGGTTCAA	3500
GCGATTCTCC TGCCTCAGCC ACCTGAGTAG CTGGGAGTAC AGGCATGTGG	3550
CACCATGCCT GGCTAATTTT TGTATTTTTA ATAGAGTTGA GATTTCACCA	3600
TGATGGTCAG GCTGGTCTCG AACTCCTGAC CTCGTGATCC GCCCGCCTCG	3650
GCCCCCAAA GTGCTGGGAT TAACGGGTGT GAGCCACGGC GCCCAGCCCA	3700
GGAAGAGTTT TTAAATTAGA GCTCTGTTTA ATTATACCAC TGGGAAATCA	3750
TGGTTACGCT TCAGGCATAT TCTTCCCCAG AGTACTACTT ACATTTTAAA	3800
TITCATITIG TAAAGITAAA TGTCAGCATT CCCTTTAAAA GTGTCCATTG	3850
TTCTTTGAAA GTAGACGTTT CAGTCATTCT TTTCAAACAA GTGTTTGTGT	3900

FIG. 17D

4200 4250 4100 4000 3950 4050 GTATTTTGT GATCTGTAAT GAAAAGAATC TGTACTGCAA GTAAAACCTA CTCCCCAAAA ATGTGTGT TTGGGTCTGC ATTAAACGCT GTAGTCCATG ACCTITIGOC AAGCIGIGG CATCGIGTG GAGTACAGGG IGCICAGCIC CAGATTGAC CTTGATTGAC TGTCAGGCAT GGCTTTGTTT CTAGTTTCAA ITCCACCGTC ATITIGAATT GITCACATGG GTAATTGGTC ATGGAAATGA AATATGAGCT ACTGCATGTA ATTCTTAAAC TGGGCTTGTC ACATTGTAT) CTGTTCTCG TTCCTTGTAC CGGATTATTC TACTCCTGCA ATGAACCCTG TGACACCGG ATTTAGCTCT TGTCGGCCTT CGTGGGGAGC TGTTTGTGTT TCATGCC

35/39

FIG. 17E

50 100 150 200 250 300 ALEKRKLFAC EEHPSHKAVW NVLGNLSEIQ GEVLSFDGNR TDKNYIRLEE IALLMGVNNN ETCRHLSCVL SGLIADLDAL DVCGRTEIRN YRREVVEDIN KLLKYLDLEE EADTTKAFDL RONHSILKIE KVLKRMREIK NELLQAQNPS HTLTKICAVQ EIIEDCMKKQ PSLPLSEDAH PSVAKINFVM CEVNKARGVL LLTKOLLALD AVDPQGEEKC KAARKQAVRL AQNILSYLDL KSDEWEY MDMGNQHPSI SRLQEIQKEV KSVEQQVIGF SGLSDDKNYK KLERILTKOL FEIDSVDTEG KGDIQQARKR AAQETERLLK ELEQNANHPH RIEIQNIFEE AQSLVREKIV PFYNGGNCVT DEFEEGIODI ILRLTHVKTG GKISLRKARY ELYLSSKTEL QGLIGQLDEV SLEKNPCIRE ARRRAVIEVQ TLITYIDLKE

17-	
F.9.)

					36	/39					
8	1.60	270	360	450	240	630	720	910	100	940	1080
TCAGACCTAG	CCTCAGCCAC	CCAACATCCT Q H P	TGACAAGAAT D K N	GCAAGCTAGG Q A R	GAACATTTTT N I F	AGGCATCCAA G I Q	AATCTGTGCG I C A	AATCAACTTC I N F	CTTATCCTGT L S C	AGTAGAAGAT V E D	CATTTTAAAA I L K
פנופננננפנ	CACCTCGCCA	GGGACGCCAA GACCGCATCC AATTCAGACT TCTTTGGTG CTTGTGAAAC TGAACACAAC AAAAGTATGG ATATGGGAAA CCAACATCCT M D M G N Q H P	TCTATTAGTA GGCTTCAGGA AATCCAAAAG GAAGTAAAAA GTGTAGAACA GCAAGTTATC GGCTTCAGTG GTCTGTCAGA TGACAAGAAT SISR LOEEIORKEVKSVEOQ ON IGFSG LSDDKN	TACAAGAAAC TGGAGAGGAT TCTAACAAAA CAGCTTTTTG AAATAGACTC TGTAGATACT GAAGGAAAAG GAGATATTCA GCAAGCTAGG Y K K L E R I L T K Q L F E I D S V D T E G K G D I Q A R	TTGAATACA E I Q	AGTTTGAAGA F E E	GATATCATTC TGAGGCTGAC ACATGTTAAA ACTGGAGGAA AAATCTCCTT GCGGAAAGCA AGGTATCAC CTTTAACCAA AATCTGTGCG D I I L R L T H V K T G G K I S L R K A R Y H T L T K I C A	CCGTTGCCAA V A K	GTGATGTGTG AGGTGAACAA GGCCCGAGGG GTCCTGATTG CACTTCTGAT GGGTGTGAAC AACAATGAGA CCTGCAGGCA CTTATCCTGT V M C E V N K A R G V L I A L L M G V N N E T C R H L S C	GTGCTCTCGG GGCTGATCGC TGACCTGGAT GCTCTAGATG TGTGCGGCCG GACAGAAATC AGAAATTATC GGAGGGAGGT AGTAGAAGAT V L S G L I A D L D A L D V C G R T E I R N Y R R E V V E D	ATCAACAAAT TATTGAAATA TCTGGATTTG GAAGAGGAAG CAGACACAAC TAAAGCATTT GACCTGAGAC AGAATCATTC CATTTTAAAA I N K L L K Y L D L E E E A D T T K A F D L R Q N H S I L K
CCCTTCACCG	GCTGATCTTC	: AAAAGTATGG M D	GCTTCAGTG 6 F S G	GAAGGAAAAG EGKG	CCACACCGGA PHRI	GTAACTGATG V T D E	AGGTATCACA R Y H T	GCACATCCTT A H P S	: AACAATGAGA N N E T	. AGAAATTATC R N Y R	GACCTGAGAC
CAGTAGCGG	CCCGTAAAGG	. TGAACACAA	GCAAGTTATC QVI	TGTAGATACT V D T	N TGCAAACCAC A N H	S AGGCAACTGC G N C	GCGGAAAGCA R K A	TTCCGAGGAT S E D	GGGTGTGAAC G V N	GACAGAAATC T E I	TAAAGCATTT K A F
CTGCTGCAGO	6676((6(A(CTTGTGAAA(GTGTAGAACA V E @	AAATAGACTO I D S	TGGAGCAGAA E a n	TTTATAATGG Y N G	AAATCTCCTT I S L	רכנדקנכננד ר ף ר	CACTTCTGAT	TGTGCGGCCG	CAGACACAAC
CCC66A6C66	GCCCAGCTCC	7.011116616	GAAGTAAAAA E V K S	CAGCTTTTTG a L F E	CTCAAAGAGT L K E L	ATTGTGCCAT I V P F	ACTGGAGGAA T G G K	AAGCAGCCTT K a P S	GTCCTGATTG V L I A	GCTCTAGATG A L D V	GAAGAGGAAG E E E A
CCNGAAGACG	6(A6(T6666	AATTCAGACT	AATCCAAAAG I Q K	TCTAACAAAA L T K	AGAACGTCTT E R L	GAGAGAGAAA R E K	ACATGTTAAA H V K	CTGCATGAAA C M K	GGCCGAGGG A R G	TGACCTGGAT D L D	TCTGGATTTG L D L
נננננננננ	TGCGAGGCAT	GACCGCATCC	GGCTTCAGGA L Q E	TGGAGAGGAT E R I	CACAGGAGAC Q E T	AGTCCCTCGT S L V	TGAGGCTGAC R L T	TAATCGAAGA I E D	AGGTGAACAA V N K	GGCTGATCGC L I A	TATTGAAATA L K Y
ונננננננננ	T(666A6666	666ACGCCAA	TCTATTAGTA S I S R	TACAAGAAAC Y K K L	AAGCGGCAG K R A A	GAGGAAGCCC E E A A	GATATCATTC D I I L	GTGCAAGAGA V Q E I	6T6AT6T6T6 V M C E	6T6CTCTC66 V L S 6	ATCAACAAAT I N K L

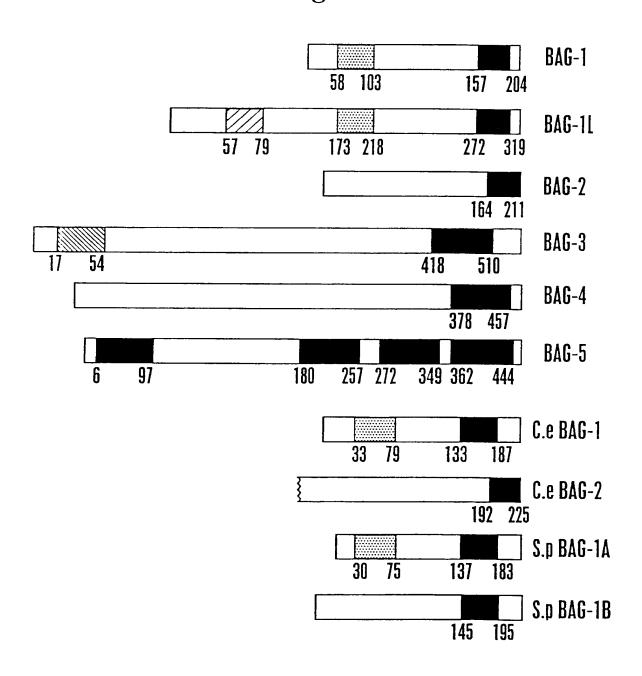
	1170	1260	1350	1440	1,260	1350	0 h h (1530	1,620	1710	1,690 1,980 2070
	CTCCAAACA S K T	AGTGATCGAG V I E	CCATAAAGCC H K A	CATCCGGCTG I R L	GAATTGCAGG GTTTAATTGG ACAGTTGGAT GAGGTAAGTC TTGAAAAAA CCCCTGCATC CGGGAAGCCA GGAGAAGAGC AGTGATCGAG E L Q G L I G Q L D E V S L E K N P C I R E A R R A V I E	CCATAAAGCC H K A	CATCCGGCTG I R L	ACAAGCTGTG Q A V	TGATACTGTT	TGATTGAAGC TATCTGTATG	TAAATGGGTT GGGAAATTTT TATAAGGCCT
	TGTACCTGAG Y L S	GGAGAAGAGC R R A	AGCACCCATC H P S	ATAAGAACTA K N Y	GGAGAAGAGC R R A	AGCACCCATC H P S	ATAAGAACTA K N Y	CTGCCAGGAA A R K	ATCTCACTTT	TCAGTATTTA TTCGGATCAT	TTGTAGATTT TATGAGAAAG TTTTTTCACT
	CCTTCTGAAT P S E L	CGGGAAGCCA R E A R	GCTTGTGAGG A C E E	AATCGAACCG N R T D	CGGGAAGCCA R E A R	GCTTGTGAGG A C E E	AATCGAACCG N R T D	AAGTGTAAGG K C K A	AATACCAGAG	TATTTCAGTC AACTTTTCCA	GAGGCAGCTT CTAGGAAAAT TAGAAGGCTT
)	AGCACAAAAC A Q N	CCCCTGCATC P C I	AAAGCTGTTT K L F	ATTTGATGGA F D G	CCCCTGCATC P C I	AAAGCTGTTT K L F	ATTTGATGGA F D G	GGGAGAAGAG G E E	GGAGTACTGA E Y	TATACGTGCA ACAGCACGTT	AAACAAATA TTACATAATA TTATGCTCGA
ָ - -	AACTTCTCCA L L Q	TTGAAAAAA E K N	TTGAGAAAAG E K R	AAGTTCTTTC V L S	TTGAAAAAA E K N	TTGAGAAAAG E K R	AAGTTCTTTC V L S	TTGATCCGCA D P Q	CTGATGAATG D E W	TTCATTGATT AAATATCATT	TTTAATCAGA AGGCATAACC TTCTGTTTCA
	ATAAAAATG I K N E	GAGGTAAGTC E V S L	AAGGAGGCCC K E A L	ATCCAGGGAG I Q G E	GAGGTAAGTC E V S L	AAGGAGGCCC K E A L	ATCCAGGGAG I Q G E	CTGGATGCTG L D A V	GACCTGAAAT D L K S	GAGCTTTCAG GTTGCAAGAC	TTTTTGCGTT ATCTAGAACT GCAGTACATG
	AATGAGAGAA M R E	ACAGTTGGAT Q L D	TATTGACTTG I D L	CTTGTCTGAG L S E	ACAGTTGGAT Q L D	TATTGACTTG I D L	CTTGTCTGAG L S E	GCTGCTAGCC L L A	CAGCTATCTC S Y L	TATGTATAGA TGCTTTTGAT	TTTGTCCTTT GTCTTTCAGA TCAAACACAA
	ATAGAAAAGG TCCTCAAGAG AATGAGAGAA ATAAAAAATG AACTTCTCCA AGCACAAAAC CCTTCTGAAT TGTACCTGAG CTCCAAAACA I E K V L K R M R E I K N E L L Q A Q N P S E L Y L S S K T	GAATTGCAGG GTTTAATTGG ACAGTTGGAT GAGGTAAGTC TTGAAAAAA CCCCTGCATC CGGGAAGCCA GGAGAAGAG AGTGATCGAG E L Q G L I G Q L D E V S L E K N P C I R E A R R A V I E	TGATCACATA I T Y	GTCTGGAACG TCCTTGGAAA CTTGTCTGAG ATCCAGGGAG AAGTTCTTTC ATTTGATGGA AATCGAACCG ATAAGAACTA CATCCGGCTG V W N V L G N L S E I Q G E V L S F D G N R T D K N Y I R L	GTTTAATTGG L I G	TGATCACATA I T Y	TCCTTGGAAA L G N	TCACCAAGCA T K a	AGGCTTGCGC AGAATATTCT CAGCTATCTC GACCTGAAAT CTGATGAATG GGAGTACTGA AATACCAGAG ATCTCACTTT TGATACTGTT R L A Q N I L S Y L D L K S D E W E Y	TATGTGCTTC CAGTATCTGC	ATGTGGTGTG GTTTGTTTGG TTTGTCCTTT TTTTGCGTT TTTAATCAGA AAACAAAATA GAGGCAGCTT TTGTAGATTT TAAATGGGTT GTGCAAGCAT TAAAATGCAG GTCTTTCAGA ATCTAGAACT AGGCATAACC TTACATAATA CTAGGAAAAT TATGAGAAAG GGGAAATTTT TGGTTAAATA AGAGTAAGGT TCAAACACAA GCAGTACATG TTCTGTTTCA TTATGCTCGA TAGAAGGCTT TTTTTTCACT TATAAGGCCT
	ATAGAAAAGG I E K V	GAATTGCAGG E L Q G	GTGCAAACTC V Q T L	GTCTGGAACG V W N V	GAATTGCAGG E L A G	GTGCAAACTC V Q T L	GTCTGGAACG V U N V	GAAGAGCTGC E E L L	AGGCTTGCGC R L A A	TTGCACTTCA AAATTCTATT	ATGTGGTGTG GTGCAAGCAT TGGTTAAATA

GATTGGTCCT	ACCCAGCTTA	999	611111161	TTGTTCAGAC	AGTCTGTTCT	TTTGTAAACA	TTTTTAGTTG	GAAAACAGC	2360
ATCTGCATTT TCCCCATCCT CTACGTT1	TCCCCATCCT	CTACGTTTTA	GAGAGGAATC	T16TTTT16T	GTGCAACATA	AGAAAATTAT	GAAAACTAAT	AGCCAAAAA	2250
CCTTTGAGAT TGCATTAAAG AGAAGGGA	TGCATTAAAG		AAGGACCAGC	AATAATACCT	TGTAAGTTGC	TTTTGTTTGT	AAAATCTGAG	CTTATAGTTT	2340
TCCTTAGTGA GTAAATTCAT	GTAAATTCAT	AAGGATGGGA	ACATTTAAAT	TAAGTTAATG	GGCCTTTAAA	AAAAAAAAG	GAAACACTCA	TACCTGTAGT	2430
TGGAGGATGA ATACTGGAGA CGGGTTA(ATACTGGAGA	CGGGTTA	ATGTCAGGTT	ATACTAAAAC	TAAATCAGAA	AGTCTGAATG	TAGCACATAA	16611(1(11	2520
CTGTTGTCCA AGGCTGTAAA	AGGCTGTAAA	ATGGACAGCC	TTGTCACACC	10000000	TGTTTTACAA	CGTGAGGGTA	GACGCTGTCA	GTAACCCAGA	2610
GGGACCAGGC	CTTCCTAGGT	TTTCTAGGCA	CTCAGCTGTT	AACCACTCAC	TTAGTAAATG	TCATAACTAC	ACCTGCTCCA	GGACCAATCA	2700
GTGAAACCTG CTCGGAATTA	CTCGGAATTA	CTCGGAATTA AAGGCTTCCT	CT666T6CCT	GCTGAACAAC	TGAGCTCATG	TCATGGGCAT	6766766777	CTCTGTTGCC	2790
TGAAAGAGCC	ATTAAAGTCA	67(676(676	AAGCATCTCT	CTTCTAAAGG	ATGTGTATTT	CCATAAATGC	TTTCTGAGGA	TCCGGTACAA	2680
	CAAAGTTCTG	AAGTGCCTTG	AGAACATGTG	6670064676	TTATAACAGA	(T(CTCCCC)	GGGTCACCTT	116((1661(2970
ATCCTGTTAG AGTACATCTT	AGTACATCTT	TGGAAATCCA	GGGTAATATT	CTCTTTCAGA	GATGCTCATT	GTGTAACTCT	GTGTAGGGAG	ATAGTCACTT	3060
TAAACAGCTC	AAAGTAGCTA		TAGCCTTAAA	TACCTAAAAG	ATGACAGAAG	CATAGCCCTT	AACAAATCTT	CAGCTTGTCT	31.50
CTCAGTATTT	CCCAATCATG	AAAATCCCTT	GCTATGTCTT	TCCTACTAGA	AATGTTCTAG	AATCGCTGGA	0.667666670	AGAGGGCAGT	3240
CGGTATTTAG GCCGTGAGCT	GCCGTGAGCT	TCCCATACTA	CTGCAGGTCC	AACTCCTGGC	AACCGCGGC	TCAAGGCAGG	TCATTGGAAT	CCACGTTTTG	3330
GCCACAGTAG TTGTAGGATT	TTGTAGGATT		ATCATAATTT	TAGAATGCTC	TTAAAATCTT	GAGGAAGAGT	TTTTATTT	TATTTATTT	3450
TGAGATGGAG	1010101		GTGCAGTGGT	GCCATCTCAG	CTCACTGCAA	CCTCCACCTC	CCAGGTTCAA	GCGATTCTCC	3510
TGCCTCAGCC	ACCTGAGTAG	CTGGGAG	AGGCATGTGG	CACCATGCCT	GGCTAATTTT	TGTATTTTA	ATAGAGTTGA	GATTTCACCA	3600
TGATGGTCAG	TGATGGTCAG GCTGGTCTCG AACTCCT	AACTCCTGAC	CTCGTGATCC	9)1))9)))9	GCCCCCCAAA	GTGCTGGGAT	TAACGGGTGT	GAGCCACGGC	38-10
GCCCAGCCCA	GCCCAGCCCA GGAAGAGTTT	TTAAATTAGA	GCTCTGTTTA	ATTATACCAC	TGGGAAATCA	TGGTTACGCT	TCAGGCATAT	TCTTCCCCAG	3780
AGTACTACTT	ACATTTTAAA	TTTCATTTTG	TAAAGTTAAA	TGTCAGCATT	CCCTTTAAAA	GTGTCCATTG	TTCTTTGAAA	GTAGACGTTT	3870
CAGTCATTCT	TTTCAAACAA GTGTTTG	GTGTTTGTGT	ACCTTTTGCC	AAGCTGTGGG	CATCGTGTGT	GAGTACAGGG	TGCTCAGCTC	TTCCACCGTC	3460
ATTTGAATT	GTTCACATGG GTAATTGG	GTAATTGGTC	ATGGAAATGA	TCAGATTGAC	CTTGATTGAC	TGTCAGGCAT	66(1116111	CTAGTTTCAA	4050
707677076	TTCCTTGTAC	CGGATTATTC	TACTCCTGCA	ATGAACCCTG	TTGACACCGG	ATTTAGCTCT	161(66((11	CGTGGGGAGC	4]40
TGTTTGTGTT	AATATGAGCT	ACTGCATGTA	ATTCTTAAAC	1666(1161(ACATTGTATT	GTATTTTGT	GATCTGTAAT	GAAAGAATC	4230
TGTACTGCAA	GTAAAACCTA	CTCCCCAAAA	ATGTGTGGCT	1166670760	ATTAAACGCT	GTAGTCCATG	TTCATGCC		4320

Fig. 17H

39/39

Fig. 18



Ubiquitin-Like BAG Domain

WW Nuclear Localization Signal

SEQUENCE LISTING

<110> Reed, John C. Takayama, Shinichi The Burnham Institute <120> Novel BAG Proteins and Nucleic Acid Molecules Encoding Them <130> FP-LJ 3646 <140> <141> <150> 09/150,489 <151> 1998-09-09 <160> 24 <170> PatentIn Ver. 2.0 <210> 1 <211> 1291 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (46)..(1080) <400> 1 acgccgcgct cagcttccat cgctgggcgg tcaacaagtg cgggc ctg gct cag cgc 57 Leu Ala Gln Arg 1 ggg ggg gcg cgg aga ccg cga ggc gac cgg gag cgg ctg ggt tcc cgg Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg Leu Gly Ser Arg 10 15 20 ctg cgc gcc ctt cgg cca ggc cgg gag ccg cgc cag tcg gag ccc ccg 153 Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln Ser Glu Pro Pro 25 gcc cag cgt ggt ccg cct ccc tct cgg cgt cca cct gcc cgg agt act 201 Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro Ala Arg Ser Thr 40 45 50 ged agd ggg dat gad dga ded add ggd gdd gdd gdd gdt dgd 249

Ala	Ser	Gly 55	His	Asp	Arg	Pro	Thr 60	Arg	Gly	Ala	Ala	Ala 65	Gly	Ala	Arg	
				aag Lys												297
				agc Ser												345
				acc Thr 105												393
				gag Glu												441
				gag Glu			-	_	-						-	489
				gag Glu												537
				gtt Val												585
		-		cag Gln 185			,					_				633
				gaa Glu												681
				tta Leu						_		_				729
				ttg Leu												777
gac	cag	ctg	gaa	gag	ttg	aat	aaa	gag	ctt	act	gga	atc	cag	cag	ggt	825

Asp (245	Gln	Leu	Glu	Glu	Leu 250	Asn	Lys	Glu	Leu	Thr 255	Gly	Ile	Gln	Gln	Gly 260	
ttt Phe	ctg Leu	ccc Pro	aag Lys	gat Asp 265	ttg Leu	caa Gln	gct Ala	gaa Glu	gct Ala 270	ctc Leu	tgc Cys	aaa Lys	ctt Leu	gat Asp 275	agg :	373
aga Arg	gta Val	aaa Lys	gcc Ala 280	aca Thr	ata Ile	gag Glu	cag Gln	ttt Phe 285	atg Met	aag Lys	atc Ile	ttg Leu	gag Glu 290	gag Glu	att Ile	921
gac Asp	aca Thr	ctg Leu 295	Ile	ctg Leu	cca Pro	gaa Glu	aat Asn 300	Phe	aaa Lys	gac Asp	agt Ser	aga Arg 305	ttg Leu	aaa Lys	agg Arg	969
aaa Lys	ggc Gly 310	Leu	g gta 1 Val	aa a Lys	aag Lys	gtt Val 315	Glr	g gca n Ala	ttc Phe	cta Leu	gcc Ala 320	Glu	tgt Cys	gac Asp	aca Thr	1017
gtg Val 325	Glu	caq Gli	g aac n Asr	ato	tgc Cys 330	s Glr	gaq Glu	g act u Thi	gaç Glu	g egg 1 Arg 335	g Let	caç Glr	j tot Ser	aca Thi	a aac Asn 340	1065
			g gco u Ala		u	aggto	gtag	caga	аааа	agg (ctgto	gctgo	ee et	tgaa	gaatg	1120
gcġ	gcca	ccag	ctc	tgcc	gtc	tctg	gatc	gg a	attt	acct	g at	ttct	tcag	ggc	tgctggg	1180
ggo	caac	tggc	cat	ttgc	caa	tttt	ccta	.ct c	tcac	actg	g tt	ctca	atga	aaa	atagtgt	1240
ct [.]	ttgt	gatt	tga	gtaa	agc	tcct	atto	tg t	ttt.	caca	a aa	aaaa	aaaa	a		1291
	10>															

<212> PRT

<213> Homo sapiens

<400> 2

Leu Ala Gln Arg Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg 10 5 1

Leu Gly Ser Arg Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln 25 20

Ser Glu Pro Pro Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro 40 35

Ala Arg Ser Thr Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala Ala Gly Ala Arg Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser Thr Arg Ser Glu Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu Glu Ala Thr Trp Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln Gly Glu Glu Met Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser Thr Arg Ser Glu Glu Val Thr Arg Glu Glu Met Ala Ala Gly Leu Thr Val Thr Val Thr His Ser Asn Glu Lys His Asp Leu His Val Thr Ser Gln Gln Gly Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val Val Glu Glu Val Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe Lys Gly Lys Ser Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly 200 205 Ile Gln Asp Gly Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro Gln Glu Glu Val Glu Leu Lys Lys Leu Lys His Leu Glu Lys Ser Val 230 235 240 Glu Lys Ile Ala Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly Ile Gln Gln Gly Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys Lys Leu Asp Arg Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile Leu Glu Glu Ile Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser

Arg Leu Lys Arg Lys Gly Leu Val Lys Lys Val Gln Ala Phe Leu Ala

305 310 315 Glu Cys Asp Thr Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu 325 330 Gln Ser Thr Asn Phe Ala Leu Ala Glu 340 345 <210> 3 <211> 1179 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (160)..(792) <400> 3 gcagccgcgg tgtcgcgaag tcctcccggg ttgcccccgc ggcgtcagag ggagggcggg 60 egeegegttg gtgaeggega eeetgeagee caaggagege teeacteget geegeeggag 120 ggccggtgac ctcttggcta ccccgcgtcg gaggcttag atg gct cag gcg aag Met Ala Gln Ala Lys 1 atc aac get aaa gee aac gag ggg ege tte tge ege tee tee atg 222 Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys Arg Ser Ser Met 10 15 20 gct gac egc tee age ege etg etg gag age etg gac eag etg gag ete 270 Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu Asp Gln Leu Glu Leu 25 agg gtt gaa gct ttg aga gaa gca gca act gct gtt gag caa gag aaa 318 Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala Val Glu Gln Glu Lys 40 45 gaa atc ctt ctg gaa atg atc cac agt atc caa aat agc cag gac atg Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln Asn Ser Gln Asp Met 55 60 65 agg cag atc agt gac gga gaa aga gaa tta aat ctg act gca aac Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu Asn Leu Thr Ala Asn 70 75

_	-	_	2 2	_			acc Thr	-	-			-	-			462
-			-	-		-	tcc Ser		_		_					510
_			-		-		ctg Leu 125	_	-	-			-	-	-	558
		_	-			-	gca Ala	_								60 b
							tcc Ser					_	_		-	654
-	-	-			-	_	aga Arg		, ,		_		_			702
			-	-	-		aag Lys							•	•	750
_				_			aat Asn 205	_	-	_	-					792
tagt	ctt	caa a	accta	aaga	gc at	ttad	cacaa	a tad	cacaa	aggt	gtaa	aaaat	iga '	taaaa	atacta	852
															gttgtt	
									-	·	-				agcaat	
															atctag	
															agctag :gggca	
			tagaa				-	- cy	geac	yyaa	cca	Jyca	<i>y</i>	١	-yyyca	1179
	- 4	- u	caya	<u> </u>		Lact	34									11/2

<210> 4

<211> 211

<212> PRT

<213> Homo sapiens

<400> 4

Met Ala Gln Ala Lys Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys

1 5 10 15

Arg Ser Ser Ser Met Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu
20 25 30

Asp Gln Leu Glu Leu Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala 35 40 45

Val Glu Gln Glu Lys Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln
50 55 60

Asn Ser Gln Asp Met Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu 65 70 75 80

Asn Leu Thr Ala Asn Arg Leu Met Gly Arg Thr Leu Thr Val Glu Val 85 90 95

Ser Val Glu Thr Ile Arg Asn Pro Gln Gln Gln Glu Ser Leu Lys His
100 105 110

Ala Thr Arg Ile Ile Asp Glu Val Val Asn Lys Phe Leu Asp Asp Leu 115 120 125

Gly Asn Ala Lys Ser His Leu Met Ser Leu Tyr Ser Ala Cys Ser Ser 130 135 140

Glu Val Pro His Gly Pro Val Asp Gln Lys Phe Gln Ser Ile Val Ile 145 150 155 160

Gly Cys Ala Leu Glu Asp Gln Lys Lys Ile Lys Arg Arg Leu Glu Thr 165 170 175

Leu Leu Arg Asn Ile Glu Asn Ser Asp Lys Ala Ile Lys Leu Leu Glu 180 185 190

His Ser Lys Gly Ala Gly Ser Lys Thr Leu Gln Gln Asn Ala Glu Ser 195 200 205

Arg Phe Asn 210

<210> 5 <211: 2528 <212: DNA <213: Homo sapiens <220 H <221 - CDS <222> (1)..(2031) <400 - 5 gcg gag etc ege atc caa eec egg gee geg gee aac tte tet gga etg Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Ala Asn Phe Ser Gly Leu 10 gac cag aag ttt cta gcc ggc cag ttg cta cct ccc ttt atc tcc tcc 96 Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser 20 25 ttc ccc tct ggc agc gag gct att tcc aga cac ttc cac ccc tct 144 Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser 35 40 ctg gcc acg tca ccc ccg cct tta att cat aaa ggt gcc cgg cgc cgg 192 Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg 50 ctt ccc gga cac gtc ggc ggc gga gag ggg ccc acg gcg gcc cgg 240 Leu Pro Gly His Val Gly Gly Gly Glu Gly Pro Thr Ala Ala Ala Arg 65 70 cca gag act egg ege eeg gag eea geg eee ege ace ege gee eea geg 288 Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala 85 90 gge aga ded caa ded age atg age ged add cad teg ded atg atg Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met 100 105 110 cag gtg gcg tcc ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp 125 120 115 gag atc aag atc gac ccg cag acc ggc tgg ccc ttc ttc gtg gac cac 432 Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His 135 aac ago ogo aco act acg tgg aac gao oog ogo gtg ooc tot gag ggo

Asn 145	Ser	Arg	Thr	Thr	Thr 150	Trp	Asn	Asp	Pro	Arg 155	Val	Pro	Ser	Glu	Gly 160	
	aag Lys						-									528
	ctg Leu															576
	ggc Gly									-		-				624
	g gtg Val 210					=			-			_	-	-		672
_	act Thr			-	-		-		_			_			-	720
	g ggc g Gly								-			-	_	_		768
	a gcg a Ala															816
	cag Gln			_	-		-	-						-	-	864
	ctg Leu 290						-	_	_		•		_		•	912
	g ggg Gly															960
	a gca o Ala															1008
gcç	g cag	agg	ggt	gag	tac	cag	acc	cac	cag	cct	gtg	tac	cac	aag	atc	1056

Ala	Gln	Arg	Gly 340	Glu	Tyr	Gln	Thr	His 345	Gln	Pro	Val	Tyr	His 350	Lys	Ile	
-		-	-		gag Glu				-			-				1104
			_	-	ggt Gly	-	_	-		_	-					1152
-	-	-			cac His 390			-			-					1200
-	_			_	cag Gln		-			-	-		_		-	1248
	_		-		aaa Lys		-	_	-				-			1296
					cac His											1344
				-	tcc Ser	-	_							_	=	1392
			-		cct Pro 470	-		-		-				-		1440
				_	ccc Pro				_	-	-	_				1488
				_	act Thr	-		-	_	-						1536
					ccc Pro											1584
gcc	atc	ctg	gag	aag	gtg	cag	ggg	ctg	gag	cag	gct	gta	gac	aac	ttt	1632

Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe 530 535 540	
gaa ggc aag aag act gac aaa aag tac ctg atg atc gaa gag tat ttg Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu 545 550 555 560	1680
acc aaa gag ctg ctg gcc ctg gat tca gtg gac ccc gag gga cga gcc Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala 565 570 575	1728
gat gtg cgt cag gcc agg aga gac ggt gtc agg aag gtt cag acc atc Asp Val Arg Gin Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile 580 585 590	1776
ttg gaa aaa ctt gaa cag aaa gcc att gat gtc cca ggt caa gtc cag Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln 595 600 605	1824
gtc tat gaa ctc cag ccc agc aac ctt gaa gca gat cag cca ctg cag Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln 610 615 620	1872
gca atc atg gag atg ggt gcc gtg gca gca gac aag ggc aag aaa aat Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn 625 630 635 640	1920
gct gga aat gca gaa gat ccc cac aca gaa acc cag cag cca gaa gcc Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala 645 650 655	1968
aca gca gcg act tca aac ccc agc agc atg aca gac acc cct ggt Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly 660 665 670	2016
aac cca gca gca ccg tagcctctgc cctgtaaaag tcagactcgg aaccgatgtg Asn Pro Ala Ala Pro 675	2071
tgctttaggg attttagttg catgcatttc agagacttta ggtcagttgg ttttgattag	2131
ctgcttggta tgcagtactt gggtgaggca aacactataa agggctaaaa gggaaaatga	2191
tgcttttctt caatattctt actcttgtac aattaangaa gttgcttgtt gtttgagaag	2251
tttaaccccg ttgcttgttc tgcagccctg tcnacttggg cacccccacc acctgttagc	2311
tgtggttgtg cactgtettt tgtagetetg gaetggaggg gtagatgggg agteaattae	2371

coatcacata aatatgaaac atttatcaga aatgttgcca ttttaatgag atgattttct 2431
tcatctcata attaaaatac ctgactttag agagagtaaa atgtgccagg agccatagga 2491
atatctgtat gttggatgac tttaatgcta catttth 2528

<210> 6

<211 → 677

<212> PRT

<213> Homo sapiens

<400 > 6

Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Ala Asn Phe Ser Gly Leu 1 5 10 15

Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser 20 25 30

Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser 35 40 45

Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg 50 55 60

Leu Pro Gly His Val Gly Gly Glu Gly Pro Thr Ala Ala Ala Arg
65 70 75 80

Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala 85 90 95

Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met 100 105 110

Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp 115 120 125

Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His 130 135 140

Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly 145 150 155 160

Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser 165 170 175

Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg

180 185 190

Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg 195 200 205

- Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe 210 215 220
- Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu 225 230 235 240
- Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val 245 250 255
- Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg 260 265 270
- Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ser Ala 275 280 285
- Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro 290 295 300
- Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg 305 310 315 320
- Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro 325 330 335
- Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile 340 345 350
- Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe 355 360 365
- Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg 370 375 380
- Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val 385 390 395 400
- Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val 405 410 415
- Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro
 420 425 430
- Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val

Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Pro Ser Glu Lys Val 4.60 Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala Pro

<210> 7 <211> 1010 <212> DNA <213> Homo sapiens <220> <221> CDS <222> (323)..(1009) <400> 7 acgatateet gtaagaccaa gaattgeaag geeagagttt gaattettat acaaatggag 60 cgtatggtoc aacatacccc ccaggecetg gggcaaatac tgcctcatac tcaggggctt 120 attatgcacc tggttatact cagaccagtt actccacaga agttccaagt acttaccgtt 180 catetggcaa cageecaact ecagtetete gttggateta teeccageag gaetgteaag 240 actgaagcac cocctettaa ggggcaggtt ccaqqatate cqccttcaca qaaccetqqa 300 atgaccotgo occattatoo tt atg gag atg gta atc gta gtg tto cac aat Met Glu Met Val Ile Val Val Phe His Asn 1 cac ggc ega etg tac gac cac aag aaa gat geg tgg get tet eet ggt 400 His Gly Arg Leu Tyr Asp His Lys Lys Asp Ala Trp Ala Ser Pro Gly 15 gct tat gga atg ggt ggc cgt tat ccc tgg cct tca tca gcg ccc tca 448 Ala Tyr Gly Met Gly Gly Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser 35 gca cca ccc ggc aat ctc tac atg act gaa agt act tca cca tgg cct 496 Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro 45 50 age agt gge tet eee cag tea eee eet tea eee eea gte cag eee 544 Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val Gln Gln Pro 65 aag gat tot toa tac coo tat ago caa toa gat caa ago atg aac ogg 592 Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg 75 80 90 cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tcg ggg aca gtg 640 His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser Gly Thr Val 95 100 105

PCT/US99/21053 WO 00/14106

aac	aat	gat	gat	tca	gat	ctt	ttg	gat	tcc	caa	gtc	cag	tat	agt	gct	688
Asn	Asn	Asp	Asp	Ser	Asp	Leu	Leu	Asp	Ser	Gln	Val	Gin	Tyr	Ser	Ala	
			110					115					120			
gag	cct	cag	ctg	tat	ggt	aat	gcc	acc	agt	gac	cat	ccc	aac	aat	caa	736
		_	_				Ala		-							
		125		-	•		130			•		135				
gat	caa	agt	agc	agt	ctt	cct	gaa	gaa	tat	gta	cct	tca	gat	gaa	agt	784
							Glu									
- 1	140					145			_		150		•			
act	cct	cca	agt	att	aaa	aaa	atc	ata	cat	ata	cta	gag	aaq	qtc	caq	832
							Ile									
155					160	-				165			•		170	
tat	ctt	gaa	caa	gaa	ata	gaa	gaa	ttt	ata	aga	aaa	aaq	aca	gac	aaa	880
		_		~	-	_	Glu		_			_		-		
- 1 -				175					180	1	_	-		185	-	
gca	tac	taa	ctt	cta	gaa	gaa	atq	cta	acc	aaq	gaa	ctt	ttq	gaa	ctq	928
_				_	_	~	Met			_	-		-	_	-	
	- 2 -		190					195		•			200			
gat	tca	att	qaa	act	gag	aac	cag	gac	tct	qta	cqq	cag	qcc	aga	aaa	976
					-	-	Gln									
•		205			-	•	210	-			,	215		-	-	
gaq	gct	gtt	tgt	aaq	att	caq	gcc	ata	ttq	gaa	a					1010
	-	-	_	-		-	Āla		_	=						
	220		-	-		225										
<210	8 <c< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></c<>															

<211> 229

<212> PRT

<213> Homo sapiens

<400> 8

Met Glu Met Val Ile Val Val Phe His Asn His Gly Arg Leu Tyr Asp 10

His Lys Lys Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly 20

Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu 35 40 45

Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln
50 55 60

Ser Pro Pro Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro 65 70 75 80

Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser 85 90 95

Val His Gln Tyr Glu Ser Ser Gly Thr Val Asn Asn Asp Asp Ser Asp 100 105 110

Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly
115 120 125

Asn Ala Thr Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Leu 130 135 140

Pro Glu Glu Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys 145 150 155 160

Lys Ile Ile His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val 165 170 175

Glu Glu Phe Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu 180 185 190

Glu Met Leu Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly
195 200 205

Gly Gln Asp Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile 210 215 220

Gln Ala Ile Leu Glu 225

<210> 9

<211> 689

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (3)..(482)

<220>

<221> unsure

<222> (105)

<223> any amino acid

<400> 9 ga gaa ata aaa aat gaa ctt ctc caa gca caa aac cct tct gaa ttg Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu tac ctg age tee aaa aca gaa ttg cag ggt tta att gga cag ttg gat 95 Tyr Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp 20 25 30 gag gta agt ntt gaa aaa aac ccc tgc atc cgg gaa gcc agg aga aga Glu Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg 35 40 gca gtg atc gag gtg caa act ctg atc aca tat att gac ttg aag gag Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu 50 55 gcc ctt gag aaa aga aag ctg ttt gct tgt gag gag cac cca tcc cat 239 Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His 65 70 75 aaa goo gto tgg aac gto ott gga aac ttg tot gag ato cag gga gaa Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu 80 85 gtt ctt tca ttt gat gga aat cga acc gat aag aac tac atc cqg ctg 335 Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu 100 105 110 gaa gag ctg ctc acc aag cag ctg cta gcc ctg gat gct gtt gat ccg 383 Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro 115 120 cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt 431 Gln Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu 130 135 140 gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag 479 Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu 145 150 155 tac tgaaatacca gagatotoac ttttgatact gttttqcact tcatatgtqc 532 Tyr 160

ttotatgtat agagagettt eagtteattg atttataegt geatatttea gteteagtat 592 ttatgattga ageaaattet atteagtate tgetgetttt gatgttgeaa gacaaatate 652 attalageae gttaaetttt eeatteggat eaaaaaa 689

<210> 10

<211> 160

<212> PRT

<213> Homo sapiens

<400> 10

Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr

1 5 10 15

Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu
20 25 30

Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg Ala 35 40 45

Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala 50 55 60

Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys 65 70 75 80

Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val 85 90 95

Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu 100 105 110

Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln 115 120 125

Giy Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala 130 135 140

<210> 11

<211> 246

<212> DNA

<213> Caenorhabditis elegans

<400 > 11
atgt:tttcc gcctcttcgt tgaaatattt cactttcttt tccaagetttt tccccatetc 60

gacctgcttt ggtttttcga gaaaaccacg ttccaaatca gcgacatctc tcaaattgag 120
atcataggct ttttgaagat tgctcaaatt atgcttctca tattgcatga gcattttgaa 180
gcccgcgtca tcaaccaaag cattttttcc acccatcaca atgattttat cattttcttt 240
aaaatt 246

<210> 12

<211> 210

<212> PRT

<213> Caenorhabditis elegans

<400> 12

Met Lys Val Asn Val Ser Cys Ser Ser Val Gln Thr Thr Ile Asp Ile
1 5 10 15

Leu Glu Glu Asn Gln Gly Glu Asp Glu Ser Ile Leu Thr Leu Gly Gln
20 25 30

Leu Arg Asp Arg Ile Ala Thr Asp Asn Asp Val Asp Val Glu Thr Met
35 40 45

Lys Leu Leu His Arg Gly Lys Phe Leu Gln Gly Ala Asp Asp Val Ser 50 55 60

Leu Ser Thr Leu Asn Phe Lys Glu Asn Asp Lys Ile Ile Val Met Gly 65 70 75 80

Gly Lys Asn Ala Leu Val Asp Asp Ala Gly Phe Lys Met Leu Met Gln 85 90 95

Tyr Glu Lys His Asn Leu Ser Asn Leu Gln Lys Ala Tyr Asp Leu Asn 100 105 110

Leu Arg Asp Val Ala Asp Leu Glu Arg Gly Phe Leu Glu Lys Pro Lys
115 120 125

Gln Val Glu Met Gly Lys Lys Leu Glu Lys Lys Val Lys Tyr Phe Asn 130 135 140

Glu Glu Ala Glu Arg His Leu Glu Thr Leu Asp Gly Met Asn Ile Ile 145 150 155 160

Thr Glu Thr Thr Pro Glu Asn Gln Ala Lys Arg Asn Arg Glu Lys Arg 165 170 Lys Thr Leu Val Asn Gly Ile Gln Thr Leu Leu Asn Gln Asn Asp Ala 180 185 Leu Leu Arg Arg Leu Gln Glu Tyr Gln Ser Val Leu Asn Gly Asp Ile 195 200 Pro Glu 210 <210> 13 <211> 1377 <212> DNA <213> Caenorhabditis elegans <220> <221> CDS <222> (1)..(1377) <400> 13 atg cca gtc gtg aac ata cca atc aaa ata ctt ggt cag aat caa tca Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser 5 10 cat agt cga agt aac too tog tot tot gtt gac aac gat cga aat caa His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln 20 30 cca cca cag cag cca cct caa ccg caa cca caa cag caa tct cag caa Pro Pro Gln Gln Pro Pro Gln Pro Gln Gln Gln Gln Ser Gln Gln 35 40 caa tac cag cag gct cca aac gtg aat acc aat atg cat cat tcc aac 192 Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn 50 55 gga ttc tca cct aac ttc cca tct cgt agt cct att ccg gac ttt ccc Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro 65 70 75 80 agt tit toa tot ggg tio coa aac gat tot gaa tgg tot tog aat tio 288 Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe 85 90

_	-						agt Ser						-			336
		_			-		gga Gly 120	_	_	-			_			384
	_	_					aga Arg	_								432
						-	cgc Arg			-	-			-		480
				-			cca Pro			-			_	_		528
		-				-	tca Ser		-				_		•	576
							ccc Pro 200									624
			_		-		cca Pro	_	_	_			_	-	-	672
							ctt Leu									720
		_			-	-	tta Leu	-	_	-		-	_	-	• •	768
-							cca Pro			-		_		-		816
	-				-		caa Gln 280	-	-	-	-			-	-	864

																101/00) / L I U J
٠		Ile					aaa Lys										912
	aaa	290 aaa	act	act	gaa	ctc	295 gaa	ato	gaa	222	aaa	300	att	~++	cat	t c.t	960
							Glu										200
							cat His						_	-	_	-	1008
	БСС	Oly	Giu	116	325	Vai	1113	ASII	Суз	330	rne	пуѕ	Leu	GIU	335	Cys	
							gca Ala				-	_			-		1056
					_	_	gtc Val	_			-		-	-	_		1104
							act Thr 375										1152
							aag Lys									_	1200
							gaa Glu										1248
							tgc Cys									-	1296
							atg Met									_	1344
							gat Asp 455				tag						1377

<210> 14

<211> 458

<212> PRT

<213> Caenorhabditis elegans

<400> 14

Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser 1 5 10 15

His Ser Arg Ser Asn Ser Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
20 25 30

Pro Pro Gln Gln Pro Gln Pro Gln Gln Gln Gln Ser Gln Gln
35 40 45

Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn 50 55 60

Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro 65 70 75 80

Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe 85 90 95

Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn 100 105 110

Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn 115 120 125

Pro Pro Met Gln Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gln 130 135 140

Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gln Ala Pro 145 150 155 160

Pro Gln Tyr Ser Gln Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr 165 170 175

Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg
180 185 190

Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu 195 200 205

Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp 210 215 220

Ser Gly Lys Gly Asp Glu Asn Leu Glu Glu Asn Ile Ala Lys Ile Thr 225 230 235 240

Ile Gly Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly
245 250 255

Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg
260 265 270

Gly Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys 275 280 285

Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg 290 295 300

Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser 305 310 315 320

Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys
325 330 335

Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr 340 345 350

Lys Thr Val Gln Val Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys 355 360 . 365

Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met 370 380

Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met 385 390 395 400

Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe 405 410 415

Leu Lys Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
420 425 430

Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr 435 440 445

Lys Ala Asp Leu Met Asp Asp Gln Ser Glu 450 455

<210> 15

<211> 588

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221: CDS

<222: (1)(588)	
<400> 15	
atg tca gaa aag act age aca gtt aca ata cae tat gga aat eag c	ga 48
Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln A	.rg
1 5 10 15	
	-+ 0.0
ttt ccg gta gca gtc aat cta aat gag acg tta agt gaa ctg att g Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile A	
20 25 30	JP
gat tta ctt gaa acg act gag att tct gag aag aaa gtc aag ctt t	tt 144
Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu P	he
35 40 45	
tac got ggo aag ogt tta aaa gao aaa aaa goo tog tta toa aaa t	tg 192
Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys L	-
50 55 60	
ggt tta aaa aat cat agt aaa att cta tgt ata aga cca cat aag c	aa 240
Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys G	
65 70 75	80
caa cga ggt too aag gaa aaa gao aog gtt gag coo got cog aaa g	cq 288
Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys A	-
85 90 95	
gaa gcg gag aat cct gta ttt tcg cgt att tct gga gaa ata aaa g	
Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys A	la
100 105 110	
ato gat cag tat gtt gac aaa gaa ott too ooc atg tac gac aat t	ac 384
Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn T	yr
115 120 125	
gta aat aaa ccg tcg aac gat cca aag cag aaa aac aaa cag aaa c	
Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys L 130 135 140	eu
130	
atg ata agt gaa cta ctt tta caa cag ctt tta aaa ttg gat gga g	tt 480
Met Ile Ser Glu Leu Leu Gin Gln Leu Leu Lys Leu Asp Gly V	al
145 150 155 1	60
gac gta ctg ggc agc gag aaa ttg cgt ttt gaa cgg aag caa ctt g	
Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu V 165 170 175	d l
105	

tot aag ato daa aaa atg ttg gat dad gtt gad daa ada ago daa gaa 576 Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu 180 185 190

gtg gcc gca tag

Val Ala Ala

195

<210> 16

<211> 195

<212> PRT

<213> Schizosaccharomyces pombe

<400> 16

Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg 1 5 10 15

Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp 20 25 30

Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe 35 40 45

Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu 50 55 60

Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln
65 70 75 80

Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala 85 90 95

Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala 100 105 110

Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr 115 120 125

Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu 130 135 140

Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val 145 150 155 160

Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val 165 170 175

Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu 180 185 Val Ala Ala 195 <210> 17 <211> 621 <212> DNA <213> Schizosaccharomyces pombe <220> <221> CDS <222> (1)..(621) <400> 17 atg tot ttt ttt acc cag ttg tgt tct atg gat aaa aaa tat tgg atc Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile tot ota got gta ttg toa gtt act gtt ttg att agc goa tta ttg aaa 96 Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys 20 25 30 aag aga gct act gaa acc gaa gat att gtc gtt gtt cat tac gat ggc 144 Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val His Tyr Asp Gly 40 gaa aag ttg aat ttt gtg ttg cga caa cca agg ctg aat atg gtt tct 192 Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser 50 55 tac act agt ttt ctt cgt cgc gtg tgc aac gca ttt tca gta atg ccc 240 Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro 70 75 80 gac aaa gcg tct ctc aag tta aac ggg gtg acc ctc aag gat ggt tca 288 Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser 85 90 ctt tcc gac caa aat gtg caa aat gga agt gaa tta gag ctc gaa tta 336 Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu 100 105 110 ccc aaa ctg agc ccg gca atg caa caa att gaa gca tat ata gat gag 384 Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu

115 120 125

ctt caa cag gat ctc gtc cct aaa att gaa gcc ttc tgc caa tcg tct Leu Gln Gln Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser 130 135 140 ccc gct tcg gca caa gat gtt caa gat ttg cat aca cgc ctt agt gaa 480 Pro Ala Ser Ala Gln Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu 145 150 155 160 aca ttg ttg gct agg atg ata aaa tta gat gct gtt aat gtt gaa gac 528 Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp 170 gac cca gaa gct cgt ctt aaa aga aaa gaa gct att cgt tta tct caa 576 Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln 180 185 caa tat ttg agt aaa cta gat tcc acc aag aat caa aac aaa tga 621 Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys 195 200

<210> 18

<211> 206

<212> PRT

<213> Schizosaccharomyces pombe

<400> 18

Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile

1 5 10 15

Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys 20 25 30

Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val His Tyr Asp Gly
35 40 45

Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser 50 55 60

Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro 65 70 75 80

Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser 85 90 95

Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu Glu Leu

100 105 110

Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu 115 120 125

Leu Gln Gln Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser 130 135 140

Pro Ala Ser Ala Gln Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu 145 150 155 160

Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp 165 170 175

Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln 180 185 190

Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys 195 200 205

<210> 19

<211> 2534

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (307)..(2034)

<400> 19

gcggagetee gcatecaace cegggeegeg gccaacttet etggaetgga ecagaagttt 60

ctagecggee agttgetace teeetttate teeteettee eetetggeag egaggagget 120

atttccagae acttccacce etetetggee acgtcaccee egeetttaat teataaaggt 180

geceggegee ggetteeegg acaegtegge ggeggagagg ggeceaegge ggeggeeegg 240

ccagagacte ggcgcccgga gccagcgccc cgcacccgcg ccccagcggg cagaccccaa 300

cccage atg age gee gee ace cae teg ecc atg atg eag gtg geg tee 348

Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser

ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg gag atc aag atc 396 Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile
15 20 25 30

				Gly	tgg Trp				Val							444
				35					40					45		
					ccg Pro											492
1111	1111	111	50	изр	FIO	ALG	val	55	261	GIU	GIY	PIO	60 60	GIU	inr	
cca	tcc	tct	gcc	aat	ggc	cct	tcc	cgg	gag	ggc	tct	agg	ctg	ccg	cct	540
Pro	Ser	Ser 65	Ala	Asn	Gly	Pro	Ser 70	Arg	Glu	Gly	Ser	Arg 75	Leu	Pro	Pro	
gct	agg	gaa	ggc	cac	cct	gtg	tac	ccc	cag	ctc	cga	сса	ggc	tac	att	588
					Pro											
					cat His											636
95					100	010	011	1110	014	105	••• 9	0111	vai	1125	110	
					cag											684
Phe	His	Val	Tyr	Pro 115	Gln	Pro	Gly	Met	Gln 120	Arg	Phe	Arg	Thr	Glu 125	Ala	
gca	gca	gcg	gct	cct	cag	agg	tcc	cag	tca	cct	ctg	cgg	ggc	atg	cca	732
Ala	Ala	Ala	Ala 130	Pro	Gln	Arg	Ser	Gln 135	Ser	Pro	Leu	Arg	Gly 140	Met	Pro	
gaa	acc	act	Cad	CC3	gat	222	020	+ ~+	aa s	030	ata	ac.	7 07	202	222	780
					Asp										-	700
		145					150					155				
gca					gcc											828
Ala	Ala 160	Gln	Pro	Pro	Ala	Ser 165	His	Gly	Pro	Glu	Arg 170	Ser	Gln	Ser	Pro	
gct	gcc	tct	gac	tgc	tca	tcc	tca	tcc	tcc	tcg	gcc	agc	ctg	cct	tcc	876
	Ala	Ser	Asp	Cys	Ser	Ser	Ser	Ser	Ser		Ala	Ser	Leu	Pro		
175					180					185					190	
tcc	g gc	agg	agc	agc	ctg	ggc	agt	cac	cag	ctc	ccg	cgg	g gg	tac	atc	924
Ser	Gly	Arg	Ser		Leu	Gly	Ser	His		Leu	Pro	Arg	Gly		Ile	
				195					200					205		
					cac											972
Ser	Ile	Pro		Ile	His	Glu	Gln		Val	Thr	Arg	Pro		Ala	Gln	
			210					215					220			

	cac His									1020
	acc Thr							-	_	1068
	cgg Arg								-	1116
	tcg Ser						-	-		1164
	ccc Pro 290									1212
	atg Met						-		_	1260
	gaa Glu					_				1308
	cca Pro									1356
	aag Lys									1404
	cca Pro 370								_	1452
	tcc Ser									1500
	cct Pro									1548

	ccc Pro							-			-	-		_		1596
	gtg Val															1644
	gac Asp				_	-		-			_				_	1692
	gcc Ala															1740
	agg Arg 480	_			-		-	-	_			_	-			1788
	cag Gln															1836
	ccc Pro														_	1884
	ggt Gly															1932
	gat Asp				-		-	-		-	-		-	_		1980
	tca Ser 560			-	_	_		-						-	gca Ala	2028
ccg Pro 575															2084	
ttt	aagti	tgc a	atgea	attto	ca ga	agact	ttaa	a gto	cagtt	iggt	ttt	tatta	age t	tgett	iggtat	2144
gca	gtaad	ctt (gggt	ggago	gc aa	aaaca	actaa	a taa	aaag	ggct	aaaa	aagga	aaa a	atgat	gcttt	2204

tottotatat tottactoty tacaaataaa gaagttgott gttgtttgag aagtttaacc 2264

cogttgotty ttotgoagco otgtotactt gggcaccccc accacctgtt agetgtggtt 2324

gtgcactgtc ttttgtagct otggactgga ggggtagatg gggagtcaat tacccatcac 2384

ataaatatga aacatttatc agaaatgtty coattttaat gagatgattt tottcatotc 2444

ataattaaaa tacctgactt tagagagagt aaaatgtgcc aggagccata ggaatatctg 2504

tatgttggat gactttaatg ctacattttc

<210> 20

<211> 575

<212> PRT

<213> Homo sapiens

<400> 20

Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser Gly Asn 1 5 10 15

Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile Asp Pro
20 25 30

Gln Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr Thr Thr 35 40 45

Trp Asr. Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr Pro Ser
50 55 60

Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro Ala Arg 65 70 75 80

Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile Pro Ile 85 90 95

Pro Val Leu His Glu Gly Ala Glu Asn Arg Gln Val His Pro Phe His 100 105 110

Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Glu Ala Ala Ala 115 120 125

Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro Glu Thr 130 140

Thr Gin Pro Asp Lys Gin Cys Gly Gln Val Ala Ala Ala Ala Ala Ala

Gln Pro Pro Ala Ser His Gly Pro Glu Arg Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ser Ala Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val Ser Glr. Pro Glu Asn Lys 305 310 315 320 Pro Glu Ser Lys Pro Gly Pro Val Gly Pro Glu Leu Pro Pro Gly His 325 330 335 Ile Pro Ile Gln Val Ile Arg Lys Glu Val Asp Ser Lys Pro Val Ser 34C Gln Lys Pro Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro Ser Thr

Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu Ala Pro

405 410 415

Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val
420 425 430

Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp 435 440 445

Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala 450 455 460

Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg 465 470 475 480

Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln 485 490 495

Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro 500 505 510

Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly
515 520 525

Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp 530 540

Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser 545 550 555 560

Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala Pro 565 570 575

<210> 21

<211> 1966

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<2222 (43)..(1416)

<400> 21

cggtgggage ggggcgggaa gcgcttcagg gcagcggate ee atg teg gee etg 54 Met Ser Ala Leu

agg eqc teg gge tac gge ecc agt gae ggt eeg tee tac gge ege tac 102

Arg 5	Arg	Ser	Gly	Tyr	Gly 10	Pro	Ser	Asp	Gly	Pro 15	Ser	Tyr	Gly	Arg	Туг 20	
					gga Gly											150
					gaa Glu										-	198
					gcg Ala											246
					ccc Pro											294
					cag Gln 90											342
					act Thr											390
					gaa Glu											438
					cca Pro											486
					gct Ala										-	534
					cca Pro 170											582
					tgg Trp											630
gca	ccc	cct	ctt	agg	ggg	cag	gtt	сса	gga	tat	ccg	cct	tca	cag	aac	6 78

Ala	Pro	Pro	Leu 200	Arg	Gly	Gln	Val	Pro 205	Gly	Tyr	Pro	Pro	Ser 210	Gln	Asn	
							tat Tyr 220									726
							gta Val									774
							ggt Gly									822
							aat Asn									870
							ccc Pro									918
							tac Tyr 300									966
							tgc Cys					Tyr				1014
	Thr						tca Ser									1062
					Gln		tat Tyr			Ala					Pro	1110
				Glr			agt Ser		Pro					Pro		1158
			Thr					Lys					. Val		gag Glu	1206
aaç	g gto	caç	g tat	ctt	ga a	a caa	a gaa	a gtá	a gaa	a gaa	a ttt	: gta	a gga	aaa	aag	1254

Lys Val Gin Tyr Leu Glu Glu Glu Val Glu Glu Phe Val Gly Lys Lys 390 395 400	
aca gac aaa gca tac tgg ctt ctg gaa gaa atg cta acc aag gaa ctt Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu Thr Lys Glu Leu 405 410 415 420	1302
ttg gaa ctg gat tca gtt gaa act ggg ggc cag gac tct gta cgg cag Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp Ser Val Arg Gln 425 430 435	1 3 50
gec aga aaa gag get gtt tgt aag att eag gee ata etg gaa aaa tta Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu Lys Leu 440 445 450	1398
gaa aaa aaa gga tta tga aaggatttag aacaaagtgg aagcctgtta Glu Lys Lys Gly Leu 455	1446
ctaacttgac caaagaacac ttgattaggt taattaccct ctttttgaaa tgcctgttga	1506
tgacaagaag caatacatto cagottttoo tttgatttta taottgaaaa actggcaaag	1566
gaatggaaga atattttagt catgaagttg ttttcagttt tcagacgaat gaatgtaata	1626
ggaaactatg gagttaccaa tattgccaag tagactcact ccttaaaaaa tttatggata	1685
tetacaaget gettattace ageaggaggg aaacacaett cacacaacag gettateaga	1746
aacctaccag atgaaactgg atataatttg agacaaacag gatgtgtttt tttaaacatc	: 1806
tggatatett gteacatttt tgtacattgt gaetgettte aacatataet teatgtgtaa	1866
tratagetta gaetttagee ttettggaet tetgttttgt tttgttattt geagtttaea	1926
aatatagtat tattototaa aaaaaaaaaa aaaaaaaaaa	1966

<210> 22

<211> 457

<212> PRT

<213> Homo sapiens

<400> 22

Met Ser Ala Leu Arg Arg Ser Gly Tyr Gly Pro Ser Asp Gly Pro Ser 1 5 10 15

Tyr Gly Arg Tyr Tyr Gly Pro Gly Gly Gly Asp Val Pro Val His Pro

20 25 30

Pro Pro Pro Leu Tyr Pro Leu Arg Pro Glu Pro Pro Gln Pro Pro Ile 35 40 45

Ser Trp Arg Val Arg Gly Gly Gly Pro Ala Glu Thr Thr Trp Leu Gly 50 55 60

Glu Gly Gly Gly Asp Gly Tyr Tyr Pro Ser Gly Gly Ala Trp Pro 65 70 75 80

Glu Prc Gly Arg Ala Gly Gly Ser His Gln Glu Gln Pro Pro Tyr Pro
85 90 95

Ser Tyr Asn Ser Asn Tyr Trp Asn Ser Thr Ala Arg Ser Arg Ala Pro 100 105 110

Tyr Pro Ser Thr Tyr Pro Val Arg Pro Glu Leu Gln Gly Gln Ser Leu 115 120 125

Asn Ser Tyr Thr Asn Gly Ala Tyr Gly Pro Thr Tyr Pro Pro Gly Pro 130 135 140

Gly Ala Asn Thr Ala Ser Tyr Ser Gly Ala Tyr Tyr Ala Pro Gly Tyr 145 150 155 160

Thr Gln Thr Ser Tyr Ser Thr Glu Val Pro Ser Thr Tyr Arg Ser Ser 165 170 175

Gly Asn Ser Pro Thr Pro Val Ser Arg Trp Ile Tyr Pro Gln Gln Asp 180 185 190

Cys Gln Thr Glu Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro 195 200 205

Pro Ser Gln Asn Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp 210 215 220

Gly Asn Arg Ser Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu 225 230 235 240

Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly Arg Tyr Pro
245 250 255

Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr 260 265 270

Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro

275 280 285

Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln 290 295 300

Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln 305 310 315

Tyr Glu Ser Ser Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp 325 330 335

Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr 340 345 350

Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu Pro Glu Glu 355 360 365

Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile 370 375 380

His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe 385 390 395 400

Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu 405 410 415

Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp 420 425 430

Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile 435 440 445

Leu Glu Lys Leu Glu Lys Lys Gly Leu 450 455

<210> 23

<211> 4308

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (247)..(1590)

<400> 23

cccccccc cccccccc ccngaagacg cccggagcgg ctgctgcagc cagtagcggc 60

continuous getgeocogo teagacetas tegggagggs tgegaggeat geagetgggs 120	Э
goccagetes ggtgeegeae ecegtaaagg getgatette caestegesa ecteagesas 180	Э
gggacgccaa gaccgcatcc aattcagact tettttggtg ettgtgaaac tgaacacaac 240	j)
aaaagt atg gat atg gga aac caa cat cct tct att agt agg ctt cag 288 Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln 1 5 10	3
gaa atc caa aag gaa gta aaa agt gta gaa cag caa gtt atc ggc ttc 33. Glu Ile Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe 15 20 25 30	5
agt ggt ctg tca gat gac aag aat tac aag aaa ctg gag agg att cta 38. Ser Gly Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu 35 40 45	1
aca aaa cag ctt ttt gaa ata gac tct gta gat act gaa gga aaa gga 433 Thr Lys Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly 50 55 60	2
gat att cag caa gct agg aag cgg gca gca cag gag aca gaa cgt ctt 480 Asp Ile Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu 65 70 75)
ctc aaa gag ttg gag cag aat gca aac cac cca cac cgg att gaa ata 528 Leu Lys Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile 80 85 90	3
Cag aac att ttt gag gaa gcc cag tcc ctc gtg aga gag aaa att gtg 570 Gln Asn Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val 95 100 105 110	ő
cca ttt tat aat gga ggc aac tgc gta act gat gag ttt gaa gaa ggc 62. Pro Phe Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly 115 120 125	1
atc caa gat atc att ctg agg ctg aca cat gtt aaa act gga gga aaa 67: Ile Gln Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys 130 135 140	2
atc tcc ttg cgg aaa gca agg tat cac act tta acc aaa atc tgt gcg 720 Ile Ser Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala 145 150 155	Э
gtg caa gag ata atc gaa gac tgc atg aaa aag cag cct tcc ctg ccg 769 Val Gln Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro	3

	160					165					170					
		-	-	_			tcc Ser	-	-							816
							GJÀ	-	_							864
						-	agg Arg				_			-		912
							cta Leu 230									960
_						-	gta Val	-	-					_		1008
		-	_	-		-	gca Ala	_				_		-	-	1056
_	_						aaa Lys		•	-	-		_	-	_	1104
					_		ctc Leu								_	1152
							ttg Leu 310									1200
							ccc Pro									1248
							ctg Leu									1296
		-		_	-	-	ttt Phe	-	-							1344

355 360 365	
aaa goo gto tgg aac gto ott gga aac ttg tot gag ato cag gga gaa Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu 370 375 380	1392
gtt ett toa tit gat gga aat oga aco gat aag aac tac ato ogg otg Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu 385 390 395	1440
gaa gag etg ete ace aag eag etg eta gee etg gat get gtt gat eeg Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro 400 405 410	1488
cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt Gln Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu 415 420 425 430	1536
gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu 435 440 445	1584
tac tga aataccagag atotcacttt tgatactgtt ttgcacttca tatgtgcttc	1640
tatgtataga gagettteag tteattgatt tataegtgea tattteagte teagtattta	1700
tgattgaage aaattetatt cagtatetge tgettttgat gttgeaagae aaatateatt	1760
acagcacgtt aacttttcca ttcggatcat tatctgtatg atgtggtgtg gtttgtttgg	1820
tttgtccttt tttttgcgtt tttaatcaga aaacaaaata gaggcagctt ttgtagattt	1880
taaatgggtt gtgcaagcat taaaatgcag gtctttcaga atctagaact aggcataacc	1940
ttacataata ctaggaaaat tatgagaaag gggaaatttt tggttaaata agagtaaggt	2000
tcaaacacaa gcagtacatg ttotgtttoa ttatgotoga tagaaggott ttttttoact	2060
tataaggeet gattggteet acceagetta acggggtggg gtttttttgt ttgtteagae	2120
agtotgttot titgtaaaca titttagtig gaaaaacago atotgoatit tooccatoot	2180
ctacgtttta gagaggaatc ttgtttttgt gtgcaacata agaaaattat gaaaactaat	2240
agccaaaaaa cctttgagat tgcattaaag agaagggata aaggaccagc aataatacct	2300
tgtaagttgc ttttgtttgt aaaatctgag cttatagttt tccttagtga gtaaattcat	2360

aaggatggga acatttaaat taagttaatg ggcctttaaa aaaaaaaaag gaaacactca 2420 tacctgtagt tggaggatga atactggaga cgggttacca atgtcaggtt atactaaaac 2480 taaatcagaa agtotgaatg tagoacataa tggttotott otgttgtoca aggotgtaaa 2540 atggacaged tigicacade tecceggige igititacaa egigagggia gaegeigica 2600 gtaacccaga gggaccagge etteetaggt titetaggea gicagetgit aaccaeteae 2660 ttagtaaatg tcataactac acctgeteea ggaccaatea gtgaaacetg eteggaatta 2720 aaggetteet etgggtgeet getgaacaac tgageteatg teatgggeat gtggttgtt 2780 ctctgttgcc tgaaagagcc attaaagtca gtcgtgcgtg aagcatctct cttctaaagg 2840 atgtgtattt ccataaatgc tttctgagga tccggtacaa aatgatttcc caaagttctg 2900 aagtgeetig agaacatgig ggicegagig tiataacaga etecteece gggicaeeti 2960 ttgcctggtc atcctgttag agtacatctt tggaaatcca gggtaatatt ctctttcaga 3020 gatgeteatt gtgtaactet gtgtagggag atagteaett taaacagete aaagtageta 3080 gctaaaggag tagccttaaa tacctaaaag atgacagaag catagccctt aacaaatctt 3140 cagettgtet eteagtattt eccaateatg aaaateeett getatgtett teetaetaga 3200 aatgttctag aatcgctgga cggtggggtc agagggcagt cggtatttag gccgtgagct 3260 teceatacta etgeaggtee aacteetgge aaccgeggge teaaggeagg teattggaat 3320 ccacgttttg gccacagtag ttgtaggatt gcttttctgt atcataattt tagaatgctc 3380 ttaaaatctt gaggaagagt tittattitt tattiatitt tgagatggag tetetgitge 3440 ccaggetgea gtgeagtggt gecateteag eteactgeaa eetecaeete eeaggtteaa 3500 gegattetee tgeeteagee acetgagtag etgggagtae aggeatgtgg caccatgeet 3560 ggctaatttt tgtattttta atagagttga gatttcacca tgatggtcag gctggtctcg 3620 aacteetgae etegtgatee geeegeeteg geeeceeaaa gtgetgggat taaegggtgt 3680 gagocacggo goocagooca ggaagagttt ttaaattaga gototgttta attataccac 3740 tgggaaatca tggttacgct tcaggcatat tcttccccag agtactactt acattttaaa 3800

tttcattttg taaagttaaa tgtcagcatt ccctttaaaa gtgtccattg ttctttgaaa 3860 gtagacgttt cagtcattct tttcaaacaa gtgtttgtg accttttgec aagetgtggg 3920 categtgtg gagtacaggg tgctcagctc ttccaccgtc attttgaatt gttcacatgg 3980 gtaattggtc atggaaatga tcagattgac cttgattgac tgtcaggcat ggctttgttt 4040 ctagtttcaa tctgttctcg ttccttgtac cggattattc tactcctgca atgaaccctg 4100 ctgacaccgg atttagctct tgtcggcctt cgtggggagc tgtttgttt aatatgagct 4160 actgcatga attctaaac tgggcttgtc acattgtatt gtatttttgt gatctgtaat 4220 gaaaagaatc tgtactgcaa gtaaaaccta ctccccaaaa atgtgtggct ttgggtctgc 4280 attaaacgct gtagtccatg ttcatgcc

<210> 24

<211> 447

<212> PRT

<213> Homo sapiens

<400> 24

Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln Glu Ile 1 5 10 15

Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe Ser Gly
20 25 30

Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu Thr Lys
35 40 45

Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly Asp Ile 50 55 60

Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu Leu Lys
65 70 75 80

Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile Gln Asn 85 90 95

Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val Pro Phe 100 105 110

Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly Ile Gln

Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys Ile Ser Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala Val Gln Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro Leu Ser Glu Asp Ala His Pro Ser Val Ala Lys Ile Asn Phe Val Met Cys Glu Val Asn Lys Ala Arg Gly Val Leu Ile Ala Leu Leu Met Gly Val Asn Asn Asn Glu Thr Cys Arg His Leu Ser Cys Val Leu Ser Gly Leu Ile Ala Asp Leu Asp Ala Leu Asp Val Cys Gly Arg Thr Glu Ile Arg Asn Tyr Arg Arg Glu Val Val Glu Asp Ile Asn Lys Leu Leu Lys Tyr Leu Asp Leu Glu Glu Ala Asp Thr Thr Lys Ala Phe Asp Leu Arg Gln Asn His Ser Ile Leu Lys Ile Glu Lys Val Leu Lys Arg Met Arg Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr Leu 290 295 Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu Val 310 315 320 Ser Leu Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys Ala

Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val Leu

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21053

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)
This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2. X Claims Nos 1, 13, 24, 25 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: No.meaningful search could be carried out because no limitations could be placed on the sequence.
3. Claims Nos. because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)
This International Searching Authority found multiple inventions in this international application, as follows:
1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention tirst mentioned in the claims; it is covered by claims Nos.:
Remark on Protest The additional search fees were accompanied by the applicant's protest. No protest accompanied the payment of additional search fees.
The process accompanies the payment of auditional scarcil tees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21053

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
ζ	Database Genbank, National Center for Biotech. Info., Accession No. G29287, MYERS, R.M., 04 October 1996, see entire reference.	2,4
(Database Genbank, National Center for Biotech. Info., Accession No. G06974, HUDSON, T., "Whitehead Institute.MIT Center for Genome Research,'19 October 1995, see entire reference.	2,4
(Database Genseq, Derwent, Alexandria, Virginia, Accession No. V81267, OTSUKA PHARM CO LTD, 'New Bcl-2 interaction prrtein gene (Bis)- useful for elucidation of the molecular mechanism of apoptosis, and in diagnosis, prevention and treatment of diseases,' 15 December 1998 see enire reference.	2-5
X	Database, Geneseq, Derwent, Alexandria, Virginia, Accession No. T19051, MATSUBARA ET AL., "Identifying gene signatures in 3'-directed human cDNA library," 01 June 1995, see entire reference.	2,4
K	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. Q90296, LA JOLLA CANCER RES FOUN. 'Human Bcl-2-associated protein BAG-1 cDNA,'18 May 1995 see entire reference.	2-5,14

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu Glu 395

Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln Gly 405

Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala Gln 420

Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu Tyr 435

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21053

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) :07N 21/02; C07K 1/00 US CL :530/387.1, 350; 435/6, 7/1; 536/23.1 According to International Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 530/387.1, 350; 435/6, 7/1; 536/23.1 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched								
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)								
C. DOC	UMENTS CONSIDERED TO BE RELEVANT							
Category*	Citation of document, with indication, where app	propriate, of the relevant passages	Relevant to claim No.					
X	US 5,652,223 A (KOHN ET AL) 29 July 1997(29/7/97) see entire 2-5, 1 document.							
X	Database Genbank-EST, National (Accession No. AA693697, HILLIER human EST Project, '16 December 19	2						
X	Database Genbank-EST, National of Accession No. AA456862, NCI_CGAl Cancer Genome Anatomy Project (CG. August 1997, see entire reference.	P, 'National Cancer Institute,	2,4					
X Furth	er documents are listed in the continuation of Box C	. See patent family annex.						
A doc to	ectal categories of cited documents: cument defining the general state of the art which is not considered be of particular relevance	*T* later document published after the into date and not in conflict with the appl the principle or theory underlying the *X* document of particular relevance, the	ication but cited to understand invention					
"L" doc cite spe "O" doc	lier document published on or after the unternational filing date cument which may throw doubts on priority claim(s) or which is ad to establish the publication date of another citation or other icial reason (as specified)	considered novel or cannot be considered when the document is taken alone. "Y" document of particular relevance; the considered to involve an inventive combined with one or more other sucl	red to involve an inventive step c claimed invention cannot be step when the document is a documents, such combination					
P doe	cument published prior to the international filing date but later than priority date claimed	being obvious to a person skilled in t "&" document member of the same patent						
	actual completion of the international search	Date of mailing of the international sea	arch report					
24 NOVE	MBER 1999	19 JAN 2000						
Commission Box PCT	nailing address of the ISA/US ner of Patents and Trademarks 1, D.C. 20231 0. (703) 305-3230	Authorized officer SHEELA J. HUFF Telephone No. (703) 308-0196	(in. for					